12-2010

Alternative Splicing in Human Colorectal Cancer

Jae Hoon Bahn
University of Tennessee - Knoxville, jbahn@ucla.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Jae Hoon Bahn entitled "Alternative Splicing in Human Colorectal Cancer." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Comparative and Experimental Medicine.

Seung Joon Baek, Major Professor

We have read this dissertation and recommend its acceptance:

Xuemin Xu, Jay Wimalasena, Sundar Venkatachalam

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Jae Hoon Bahn entitled "Alternative Splicing in Human Colorectal Cancer." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Comparative and Experimental Medicine.

Seung Joon Baek

Major Professor

We have read this dissertation
and recommend its acceptance:

Xuemin Xu

Jay Wimalasena

Sundar Venkatachalam

Accepted for the Council:

Carolyn R. Hodges

Vice Chancellor and Dean of Graduate Studies

(Original signatures are on file with official student records)

Alternative Splicing in Human Colorectal Cancer

A Dissertation Presented for the Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Jae Hoon Bahn

December 2010

DEDICATION

This dissertation is simply impossible without my parents, Geum Hwan Bahn and Jin Myoung Che. They bore me, raised me, loved me, and supported me spiritually throughout my life. To them I dedicate this dissertation.

ACKNOLWLEDGEMENTS

It is a pleasure to thank the many people who made my dissertation possible. First and foremost, I would like to express my truthful appreciation to my advisor Dr. Seung Joon Baek for his continuous support of my Ph. D study and research, for his patience, motivation, enthusiasm and immense knowledge. I could not have imagined having a better advisor and mentor for my Ph. D study than him.

I gratefully acknowledge the members of my committee, Dr. Xuemin Xu, Dr. Jay Wimalasena and Dr. Sundar Venkatachalam for their time and helpful guidance.

I would also like to thank my lab members: Dr. Seong-ho Lee, Jason L. Liggett, Nichelle C. Whitlock, Leon Richardson, Kyungwon Min, and Thararat Nualsanit (Gib) for their friendship and help during the past three years.

My deepest gratitude goes to my parents and parents in-law who were always beside me and encouraged me to come this far. I am eternally grateful for their unconditional support and prayers.

Last but not least, I greatly express my appreciation to my wife Dr. Hye Mee Joo for her dedication, love and persistent confidence in me. I could not have done this without her.

ABSTRACT

Most human genes undergo alternative splicing, and many abnormal splicing processes are associated with human diseases. However, the molecular relationship between alternative splicing and tumorigenesis is not well understood. Here, we identified novel Krüppel-like factor 4 (KLF4) splicing variants produced by exon skipping in human cancer cell lines as well as colon tumor tissues. To elucidate the mechanism involved in KLF4 alternative splicing, we developed KLF4 minigene system and found that RNA binding motif protein 5 (RBM5) plays an important role in KLF4 splicing, as assessed by gain and loss of functional studies. Several anti-tumorigenic compounds were also tested for KLF4 splicing. Interestingly, sulindac sulfide restored wild type KLF4 ($KLF4_L$) expression and this is mediated by dephosphorylation of RBM5. Another splicing variant, small KLF4 ($KLF4_S$), localizes in the cytoplasm and nucleus, and antagonizes transcriptional activity of wild type KLF4. Our data suggest that RBM5 plays a pivotal role in the alternative splicing of KLF4, and these splicing variant forms may impact tumorigenesis.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

CHAPTER 2

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

ABBREVIATIONS

A…….…….....	Adenine
aa……...	Amino acid
A complex…….	ATP dependent spliceosomal complex
Amp……..	Ampicillin
APC…….……..	Adenomatous polyposis
ASF/SF2……	Alternative splicing factor/spicing factor 2
ATP…….....	Adenosine triphosphate
bp……..	Base pair
BPS……...	Branch point sequence
BSA……..	Bovine serum albumin
C……..……..	Cytidine
CHX…….	Cycloheximide
COX…….	Cycloxygenase
C-terminal…….	Carboxyl terminal
DAPI…….	4’6’-diamidino-2-phenylindole
dH_20…….…….	Deionized water
DMEM…..…….	Dulbecco’s modified Eagle’s medium
DMSO…….……..	Dimethyl sulfoxide
dNTP…….....	Deoxynucleotide triphosphate
DTT…….…….	Dithiothreitol

ECL…….....	Enhanced chemiluminescence
E complex…….....	Early spliceosomal complex
EDTA……	Ethylenediamine tetraactic acid
ESE……...	Exonic splicing enhancer
ESS…….……..	Exonic splicing silencer
EtBr……..	Ethidium bromide
FAP…….……..	Familial adenomatous polyposis
G……	Guanosine
hnRNP……..	Heterogenous nuclear ribonucleoprotein
HNPCC.....	Hereditary nonpolyposis colorectal cancer syndrome
ISE…….	Intronic splicing enhancer
ISS……	Iintronic splicing silencer
kb…….…….	Kilo base
kDa……	Kilo Dalton
KLF4…….……	Kruppel like factor 4
l…….……	Liter
LB…….…….	Luria broth
M……...	Molar
MMR…….	Mutation mismatch repair
mRNA……...	Messenger ribonucleic acid
NE…….	Nuclear extract
NMD…….……	Nonsense-mediated decay
NP-40…….……..	Nonidet P-40

NSAIDs ...……	Nonsteroidal ant-inflammatory drugs
NSL…….	Nuclear localization signal
N-ternimal….....	Amino terminal
ORF…….	Open reading frame
PAGE……..……..	Polyacrylamide gel electrophoresis
PBS…..……..	Phosphate-buffered saline
PCR…….……..	Polymerase chain eaction
PMSF…….……...	Phenylmethylsulfonylfluoride
PTB……..	Polypyrimidine Tract Binding protein
PTC…….……..	Premature termination codon
RBM……..	RNA binding motif
RNA……..	Ribonucleic Acid
RPMI……	Roswell Park Memorial Institute cell culture media
RRM…….....	RNA recognition motif
RS domain..……..	Arginine/serine-rich domain
RT-PCR..……...	Reverse transcriptase-polymerase chain reaction
SC35……..……	Spliceosome component of 35 kDa
SDS…….……..	Sodium dodecyl sufate
snRNP……...	Small nuclear ribonucleoprotein particle
SR…….	Serine/arginine
SS……..	Sulindac sulfide
Taq……	*Thermus aquaticus*
TAE……..	Tris/acetic acid/EDTA

TBE……..	Tris/boric acid/EDTA
TE…….	Tris/EDTA
U2AF…….	U2 snRNP auxiliary factor
UTR…….	Untranslated region
UV…….	Ultraviolet
wt……..	Wild type

CHAPTER 1

Introduction:

Alternative splicing and colorectal cancer

1.1. Colorectal cancer

Colorectal cancer is the third most common cancer and leading cause of cancer death in the United States; approximately 5-7% of individuals developing colorectal cancer. Based on the National Cancer Institute Surveillance Epidemiology and End Results (SEER) Stat Fact Sheet, it is estimated that 142,570 people (72,090 men and 70,480 women) will be diagnosed with and 51,370 people will die of colorectal cancer in 2010. The risk of colorectal cancer starts to increase before age 50 and significantly boosts in the 60s and 70s (Chan and Giovannucci 2010). Around 70% of colorectal cancers are found in the first six feet of the colon and the other 30% occur in the last 10 inches of the rectum (Figure 1-1). It is evident that further studies with molecular-based approach are required to reduce incidence and risk of colorectal cancer.

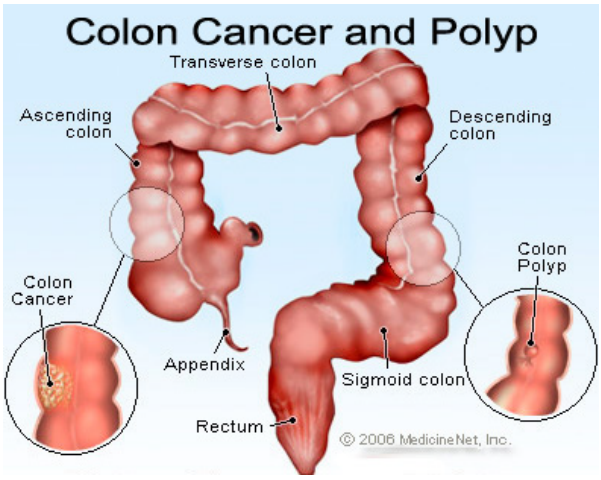

Figure 1-1. Colon cancer and polyp. The human colon has four sections, including ascending, transverse, descending, and sigmoid colon. Benign polyps develop when DNA damage occurs in the inner lining of the colon cell (Right). Over the years, polyps in the colon are able to become colon cancer (Left). Illustration from Medicine Net. Inc.

1.2. Hereditary colorectal cancer

There are many risk factors for colorectal cancer, such as genetic background, age, diet, and carcinogen exposure. About 67-95% of colorectal cancers are sporadic and the other 5-33% are hereditary (Burt and Neklason 2005, Grady 2003) (Figure 1-2).

In case of genetic factor, hereditary colorectal cancer can be classified into four groups depending on types and patterns of adenomas (Abdel-Rahman and Peltomaki 2004). The first group is more than thousands of adenomas polyps are able to become malignant including familial adenomatous polyposis (FAP) and Turcot's syndrome. The second group is less than 1000 adenomatous polyps within the colon such as MYH-associated polyposis. The third group is a few adenomas with the capacity to turn malignant including hereditary nonpolyposis colorectal cancer syndrome (HNPCC), also known as Lynch syndrome. Final group is syndromes with hamartomatous polyps which contain a mixture of normal and inflammatory polyps, such as Peutz-Jeghers syndrome, familial juvenile polyposis and Cowden syndrome (Schreibman et al 2005). However, FAP and HNPCC is the most common hereditary colorectal cancer (Figure 1-2).

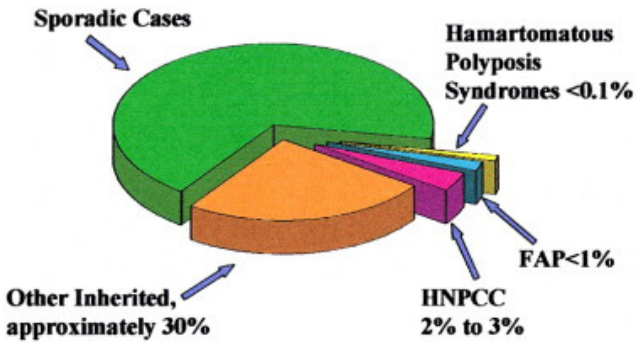

Figure 1-2. Colon cancer and inheritance. Around 33% cases of colorectal cancer is hereditary and rest of cases are sporadic (Figure from Burt and Neklason 2005).

FAP is an autosomal dominant syndrome in which over a thousand polyps develop in the patient's colon. FAP is caused by mutation of adenomatous polyposis (APC) gene, which is a tumor-suppressor gene (Burt and Neklason 2005). Mutation of APC causes constitutive activation of Wnt-signaling pathway, resulting in uncontrolled cell growth (Goss and Groden 2000). Approximately 95% of APC mutations produce truncated APC protein by frameshift or nonsense mutation (Sieber et al 2000). Because these truncations cause loss of the domains required for binding to β-catenin and to microtubules. The interaction of APC with β-catenin and microtubules is essential for its tumor suppressor activity (Aoki and Taketo 2007).

HNPCC is also an autosomal dominant syndrome which has colon cancer as well as endometrium, ovary, stomach, small intestine, gastric, urinary tract, renal cell, gallbladder, skin, and brain (Lynch et al 1993). HNPCC is associated with mutations of mutation mismatch repair (MMR) gene including MutL homolog 1 (MLH1), MutS homolog 2 (MSH2), and MutS homolog 6 (MSH6) (Burt and Neklason 2005). MMR system, which is DNA mismatch repair system, consists of a complex of proteins that recognize and repair the mismatched base-pairs during DNA replication. In addition, Some HNPCC patients have microsatellite instability (MSI), which is mutation by defect of MMR system in repeated DNA sequence within a chromosome. MSI cancer cells show 100 times more mutation rate than microsatellite stable cancer cells (Lynch 1999). Therefore, mutation in the MMR gene showed severe phenotypes by accumulations of mismatch repairs. Although there has been many studies involving hereditary colorectal cancer, sporadic colorectal cancer accounts for around 70% of colorectal cancers.

Therefore, the detailed molecular mechanism of sporadic colorectal cancer needs to be investigated.

1.3. Genetic markers in colorectal cancer

Colon cancer progression involves the accumulation of sequential events that either activate oncogenes or inhibit the action of tumor suppressor genes, resulting in the transformation of normal colonic epithelial cells into adenocarcinoma which takes between 10 to 15 years (Figure 1-3). The most commonly affected genes in this process including APC, KRAS, BRAF, SMAD4, BAX, CDC4, and p53 (Grady 2005, Walther et al 2009).

Many reports have suggested that mutations in APC gene are responsible for colorectal cancer. Generally, APC plays an important role in the regulation of cytoplasmic β-catenin by formation of multi-protein complex with Casein Kinase 1α (CK1α), Axin, and glycogen synthase kinase-3β (GSK-3β). Once β-catenin is bound to this complex, it is phosphorylated by CK1α and GSK-3β, which promotes its ubiquitin-mediated proteasomal degradation (Aoki and Taketo 2007, Fearon 1995).

On the other hand, Wnt signaling pathway involves the stabilization and nuclear translocation of β-catenin. Once Wnt ligand binds and activates their membrane receptor, Frizzled, cytoplasmic Dishevelled (Dvl) induces phosphorylation of Axin and degradation of down steam targets. As a result of the degradation, β-catenin phosphorylation is prevented and is free to translocate into the cell nucleus. Nuclear β-catenin is a co-activator of the T-Cell Factor (TCF)/Lymphoid Enhancer Factor (LEF) family of transcription factors that induce the transcription of several target genes that

control tumorigenesis including cyclin D1, CD44, PPARδ, and c-MYC (He et al 1998, Tetsu and McCormick 1999, Wielenga et al 1999).

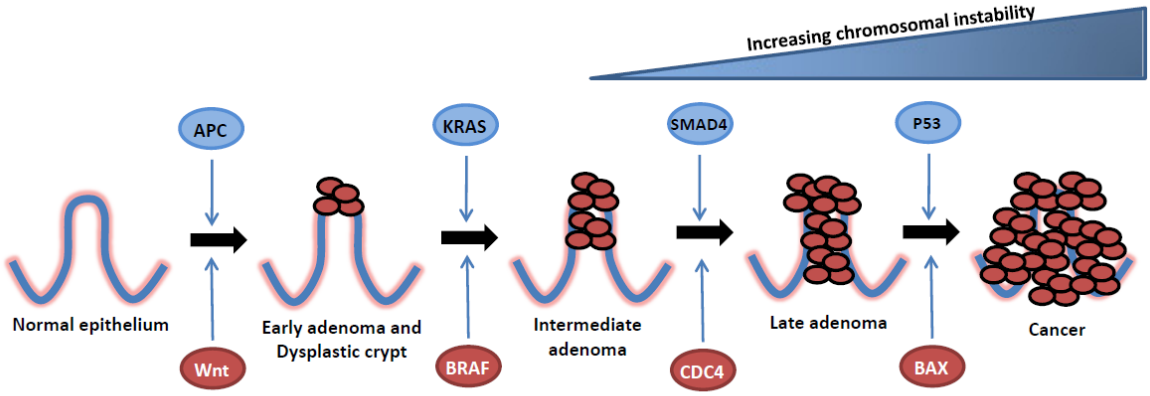

Figure 1-3. Chromosomal instability model of colorectal cancer. The initial step in tumorigenesis is associated with loss of adenomatous polyposis (APC) to form early adenoma. Mutations in KRAS, which is a small GTPase, are allowed to form intermediate adenoma. Next, late adenoma acquires loss of SMAD4, which is downstream of transforming growth factor-β (TGF-β) in chromosome 18q. Finally, mutations in p53 result in carcinoma. In addition, microsatellite instability including Wnt signaling pathway, BRAF, CDC4, and BAX is linked to each step of cancer progression (modified from Walther et al., 2009).

It has been reported that mutation of RAS in human cancer is common and around 20% of all human cancer has a mutation in one of the RAS genes. RAS is a protein that plays an important role in controlling the activity of several signaling pathways that are responsible for the regulation of cell proliferation. The presence of mutant RAS is responsible for the deregulation of cellular processes such as programmed cell death, angiogenesis, and tumor invasion. Three different members of the RAS family are found to be mutated in human cancer, including H-RAS, K-RAS and N-RAS (Downward 2003). Especially K-RAS is important in colorectal tumorigenesis. The target of RAS is RAF, a serine/threonine kinase. Once phosphorylated, RAF activates mitogen-activated protein kinase (MAPK) cascades, including MEK1/2 and Erk1/2. The activation of Erk induces its translocation to the cell nucleus, where it promotes the formation of the AP-1 transcription factor (FOS/JUN) that regulates the expression of proteins such as cyclin D1. Several reports have established that the particular effects of the RAS signaling pathway are strongly influenced by the cellular context. For example, the presence of activated RAS in immortalized cells promotes tumorigenesis, whereas in primary cells, activated RAS can induce cell cycle arrest (Mulder 2000).

Many studies show that TGF-β has the ability to regulate both cell proliferation and apoptosis at different stages of colon cancer development. SMAD is downstream of this pathway and also involved in transition stage from intermediate adenoma to late adenoma in coloreactal tumorigenesis (Derynck and Zhang 2003, Fearon and Vogelstein 1990). Especially, SMAD4 is an important tumor suppressor gene is located in chromosome 18q21, a region with frequent genetic losses in colorectal cancer (Alazzouzi et al 2005). The receptor-phosphorylated SMAD2 and SMAD3 demonstrate high affinity for SMAD4

binding. Once the complex SMAD2/3 and SMAD4 is formed, the complex translocates into the nucleus where it regulates the transcription of genes regulated by TGF-β.

1.4 KLF4 expression in cancer

So far, at least 25 members of Krüppel-like factor (KLF) family have been reported in human including SP1-like and KLF-like factors (Dang et al 2000b, Kaczynski et al 2003, Suske et al 2005, Turner and Crossley 1999). Particularly, KLF4 is strongly associated with tumorigenesis and regulation of proliferation of GI tract epithelium (Wei et al 2006). KLF4 is a member of the C2H2-type zinc finger protein family, also known as gut-enriched KLF (GKLF) or epithelial zinc finger (EZF) (Garrett-Sinha et al 1996). KLF4 binds to GC-rich elements, which is a conserved DNA binding sequence CACCC in the promoter region (Sogawa et al 1993). KLF4 acts as a transcription regulator to diverse target genes involved in cell differentiation, proliferation, and apoptosis (Katz et al 2002, Rowland and Peeper 2006, Wei et al 2006, Yoon et al 2003) (Table 1).

KLF4 protein is highly expressed in epithelial tissues, including the skin, lung, and intestine. In addition, KLF4 is detected in the differentiated terminal epithelial cells in the villus border of the small intestine and the upper crypt region of the large intestine (Ton-That et al 1997). Significant reduction of KLF4 mRNA levels was detected in colorectal adenoma and adenocarcinoma, compared with normal tissues, suggesting relevance of KLF4 in colorectal tumorigenesis (Dang et al 2000a, Wei et al 2005).

In vivo study using KLF4 knockout mice showed that KLF4 is essential for the barrier function of skin and for terminal differentiation of goblet cells in the colon of newborn mice (Katz et al 2002, Segre et al 1999). Moreover, it has been shown that KLF4

interacts with β-catenin and inhibits β-catenin signaling, suggesting that cross talk between KLF4 and β-catenin plays an important role in colorectal tumorigenesis (Zhang et al 2006).

It has been shown that KLF4 has dual functions such as tumor suppressor and oncogenic activity in tumorigenesis (Rowland and Peeper 2006, Wei et al 2006). In tumor suppressor aspect of KLF4, deletion mutation and methylation on the KLF4 gene locus were found in coloreactal and gastric cancers (Wei et al 2005, Zhao et al 2004). In addition, gastric tissue specific KLF4 knockout mice showed development of hyperplasia and polyps in stomach compared with wild type (Katz et al 2005). Investigation of human gastric cancer patients revealed that loss of KLF4 takes places at early stages in the progression of cancer (Wei et al 2005).

In contrast, ectopic overexpression of KLF4 in basal keratinocytes using inducible transgenic mice showed that squamous hyperplasia and dysplasia, indicating the oncogenic effect of KLF4 (Foster et al 2005), and overexpression of KLF4 was found around 70% of primary human breast cancers (Foster et al 2000). KLF4 has been shown as a downstream of p53 such that KLF4 is required for p53-mediated induction of cyclin-dependent kinase inhibitor 1A (CDKN1A or $p21^{CIP/WAF}$) in response to DNA damage resulting cell cycle arrest (Zhang et al 2000). Even though many clinical evidences and genetic and cell biology studies have shown a strong association between KLF4 and tumorigenesis, the regulation of KLF4 itself including transcription, post-transcription, and translation is not investigated in detail to provide a definitive role of KLF4 during tumorigenesis.

Table 1. KLF4 regulated genes (modified from Wei et al., 2006).

Gene name	Gene function	Expression level
Cyclin D1	Cell cycle	Down regulation
Orinithine decarboxylase	Cell proliferation	Down regulation
SM α-actin	Cell differentiation	Down regulation
Laminin1	Tissue homeostasis	Down regulation
Histone decarboxylase	Enzyme	Down regulation
CD11d	Cell adhesion	Down regulation
CYP1A1	Drug-metabolizing enzyme	Down regulation
p21 (WAF1/Cip1)	Cyclin-dependent kinase inhibitor	Down regulation
Keratin 4	Cell differentiation	Up regulation
Keratin 19	Stem cell marker	Up regulation
Laminin α 3A	Cell differentiation	Up regulation
Laminin γ-1 chain	Cell differentiation	Up regulation
Intestinal alkaline phosphatase	Cell differentiation	Up regulation
Small praline-rich protein 1A	Cell differentiation	Up regulation
Cytokeratin 4	Cell differentiation	Up regulation
$p27^{KIP1}$	Cell cycle	Up regulation
p53	Cell cycle, DNA damage	Up regulation
A33antigen	Intestinal epithelial cell marker	Up regulation
Urokinase-type plasminogen activator receptor	Proteolysis	Up regulation

1.5 NSAIDs in colorectal cancer

Nonsteroidal ant-inflammatory drugs (NSAIDs), including aspirin, indomethacin, piroxicam, ibuprofen, and sulindac, are used to treat pain, fever, and inflammation (Ulrich et al 2006). NSAIDs inhibit the activity of cycloxygenase-2 (COX-2) enzymes that plays a key role in the prostaglandin biosynthesis (Vane 1971). Arachidonic acid (AA), the common precursor of eicosanoids, is stored at the membrane glycerophospholipids and released by the hydrolytic action of phospholipase A2 (PLA2) enzymes. Next, AA is metabolized to the unstable intermediate prostanoid, PGH_2, by the action of COX enzymes (COX-1 and COX-2). It is generally considered that COX-1 is constitutively expressed in a variety of cells and plays a housekeeping role. However, COX-2 is a stimulus-inducible enzyme that is associated with inflammation, pain, fever and cancer. In table 2, COX-1 or COX-2 specific NSAIDs are listed (Ulrich et al 2006). In addition to their anti-inflammatory activity, NSAIDs are effective chemopreventive agents against colorectal cancer (Chan et al 2005, Thun et al 1991). However, long term treatment of aspirin and other NSAIDs can result in gastrointestinal ulceration and bleeding. This gastric problem is due to chronic inhibition of prostaglandin production via COX-1 inhibition in the gastric mucosa (Williams et al 1999). COX-2 selective inhibitors do not exhibit these side effects (Masferrer et al 1994); however, recent studies suggest that COX-2 selective inhibitors exhibit cardiovascular toxicity (McGettigan and Henry 2006, Solomon et al 2005).

It has been shown that sulindac treatment to FAP patients showed complete regression of colorectal adenomatous polyps in some cases, and partial regression (Luk 1996). Sulindac is a pro-drug metabolized by the liver and the colonic bacteria by irreversible

oxidation to sulindac sulfone and reversible reduction to sulindac sulfide (Figure 1-4). The animal studies with sulindac derivatives showed approximately twice as potent as sulindac in suppressing tissue inflammation (Hare et al 1977, Shen and Winter 1977). In addition, sulindac sulfide is at least 5,000 fold more potent than sulindac sulfone at inhibiting prostaglandin synthesis (Piazza et al 1997). However, molecular mechanism of sulindac sulfide to prevent colorectal tumorigenesis has not been elucidated in details.

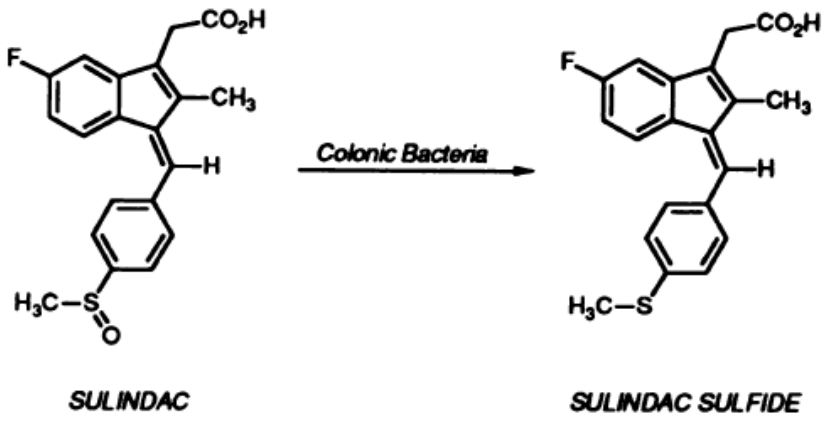

Figure 1-4. Sulindac sulfide is a metabolite from sulindac by colonic bacteria (Marnett 1992).

Table 2. NSAIDs and COX inhibition (modified from Ulrich et al., 2006).

Family/class	Drug	COX or COX2 inhibitor	Approved indications
Arylacetic acid	Naproxen	COX1 and COX2	Rheumatic diseases, inflammation, pain
Phenylalkanoic acid	Flurbiprofen	COX1 selective	Rheumatic diseases
Benzeneacetic acid	Diclofenac	COX2 selective	Rheumatic diseases, pain
Indene	Sulindac	COX1 and COX2	Rheumatic diseases, inflammation
	Indomethacin	COX1 and COX2	Rheumatic diseases. Inflammation
Propionic acid	Oxaprozin	COX1 and COX2	Rheumatic diseases
	Ibuprofen	COX1 and COX2	Rheumatic diseases, pain, fever
	Ketoprofen	COX1 selective	Rheumatic diseases, pain
Salicylate	Aspirin	COX1 and COX2	Rheumatic and heart diseases, pain, inflammation
	Disalcid	COX1 selective	Rheumatic diseases, pain
Pyranocarbonic acid	Etodolac	COX2 selective	Rheumatic diseases, pain
Pyrroleacetic acid	Tolmetin	COX1 and COX2	Rheumatic diseases
Oxicam	Piroxicam	COX1 and COX2	Rheumatic diseases
	Meloxicam	COX2 selective	Rheumatic diseases
	Lornoxicam	COX1 and COX2	Rheumatic diseases, pain
Naphthyalkanone	Nabumetone	COX1 and COX2	Rheumatic diseases
Pyrazole	Valdecoxib	COX2 selective	Rheumatic diseases, pain
	Celecoxib	COX2 selective	Rheumatic diseases, pain
Furanone	Rofecoxib	COX2 selective	Rheumatic diseases, pain

1.6 mRNA splicing

Most of eukaryote genes contain multiple exons and introns. In order to eliminate non-coding regions of precursor mRNAs (pre-mRNAs), conserved RNA sequences are marked in boundaries between exons and introns. 5' splice site (donor) and 3' splice site (acceptor) refer to as the GU-AG role. The branch point sequence (BPS) is necessary to mRNA splicing process as well (Figure 1-5). The splicing reaction is mediated by the spliceosome, and consists of five small nuclear RNA and core protein with hundreds of associated proteins (Black 2003). Most of splicing processes obey the GU-AG rule; however, some minor case reported the AU-AC rule in exon-intron recognition sites (Tarn and Steitz 1996). Depending upon nucleotides variation from the consensus sequences on the 5' or 3' splice site, the binding affinity of the spliceosome could be changed.

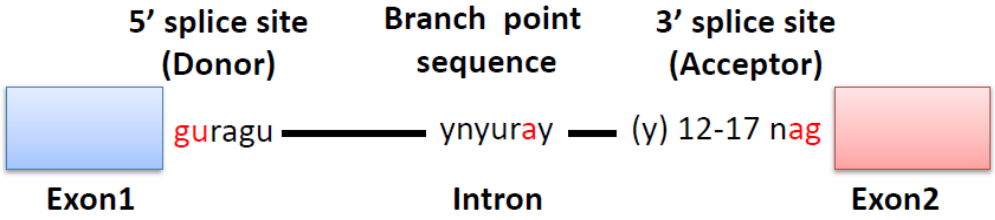

Figure 1-5. Consensus sequences for the mRNA splicing. The 5' splice site (donor) begins with a GU dinucleotide and a 3' splice site (acceptor) ends with AG dinucleotide, indicated by red letters. The branch point sequence (branch site) is located within the middle of intron. The letter "y" indicates pyrimidine and n is any nucleotide (modified from Srebrow and Kornblihtt 2006).

1.7 Biochemical reaction of mRNA splicing

There is two-step biochemical reaction in mRNA splicing via spliceosome and self-splicing transesterification (Figure 1-6). Initial biochemical step is the 2'-hydroxyl group of the branch point adenosine within intron starts nucleophilic attack at the 5' splice site and then, intron makes a lariat intermediate form. In the second biochemical reaction, 3'-hydroxyl group on the cleaved 5' splice site attacks to the phosphate at the 3' splice site and then, the two exons are ligated to each other and the intron lariat are released (Black 2003, Fedor 2008).

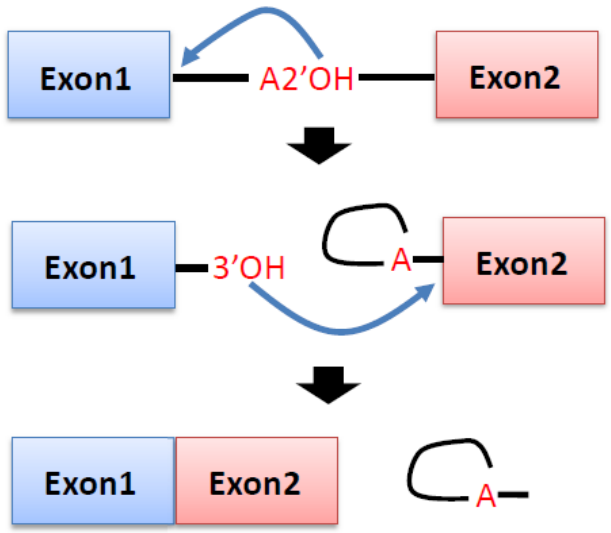

Figure 1-6. Biochemical reaction of mRNA splicing. The first step produces a intron lariat by nucleophilic attack from 2'-hydroxyl adenosine to the 5' splice site (donor). The second step produces a free intron lariat and joined two exons by the nucleophilic attack from donor to acceptor (3' splice site). Figure is modified from Fedor 2008.

1.8 Dynamic assembly of the spliceosome

The spliceosome is a major protein complex for the pre-mRNA splicing reaction which is consists of five uridine-rich small ribonucleoprotein (snRNP) complexes and over 200 accessory proteins (Black 2003, Jurica and Moore 2003). U1, U2, U4, U5, and U6 snRNP interact with the conserved exon-intron boundaries to catalyze intron excision and exon ligation (House and Lynch 2008).

A series of spliceosome components dynamically assemble to make up several intermediate complexes including the E, A, B, and C complex. For the first step, the U1 snRNP binds to the 5' splice site and the U2 auxiliary factor heterodimer (U2AF65/35) binds to the polypyrimidine tract and the 3' splice site. The U1 and U2AF bound at the end of intron becomes the early (E) complex or commitment complex by ATP-independent formation (Reed 1996). The E complex recruits the U2 snRNP, which binds to the branch point, to form the A complex. Additional binding of tri-snRNP including U4, U5, and U6 becomes the B complex by SR proteins which facilitate the formation of the B complex in an ATP-dependent manner (Roscigno and Garcia-Blanco 1995). Finally, the catalytic (C) complex is formed by the extensive rearrangement of snRNPs. For example, the U1 snRNP is replaced with the U6 snRNP to interact with the 5' splice site and the U1 and U4 snRNP is released from the spliceosome complex (Figure 1-7). The C complex catalyzes the two step biochemical reaction of mRNA splicing (Figure 1-6) as a result; exons are joined and exported to the cytoplasm for the next translation. The spliced lariat introns are degraded by several enzymes and many spliceosome components are recycled (Cheng et al 2006, Martin et al 2002).

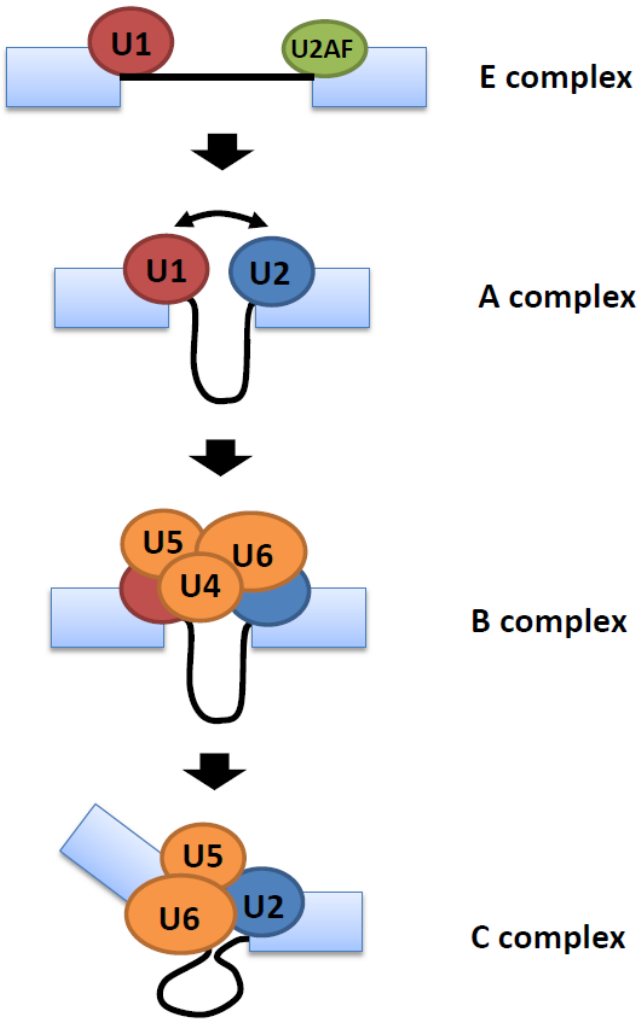

Figure 1-7. Dynamic assembly of spliceosome. For the E complex the U1 snRNP binds to the 5' splice site and U2AF binds to the polypyrimidine tract and 3' splice site. Next, the U2 snRNP binds to the branch point to form the A complex. Subsequently, the U4, U5, and U6 snRNP is recruited to the spliceosome machinery for the B complex. Finally, dynamic rearrangement takes place in the C complex; U6 snRNP replaces U1 snRNP at the 5' splice site. The U1 and U4 snRNP is released from the C complex (modified from House and Lynch 2008).

1.9 Alternative splicing patterns

Alternative splicing is a major post-transcription regulation mechanism to generate multiple gene products from single primary transcript in higher eukaryotes (Black 2003). In 1977, the first alternative splicing was reported that showed adenoviruses produce two different mRNA from the same parental gene (Berget et al 1977, Chow et al 1977a, Chow et al 1977b). More than 70 % of human genes undergo the alternative splicing (Lander et al 2001, Matlin et al 2005). In addition, the exon-exon junction microarray revealed that at least 74% of multi-exon genes undergo alternative splicing (Johnson et al 2003). Therefore, alternative splicing is the primary source of the protein diversity in higher eukaryotes including human.

There are six different modes of alternative splicing (Figure1-8). The constitutive splicing pattern is the most common splicing type which contains all exons and removes all introns. Second common type of alternative splicing is a cassette exon or exon skipping which include or exclude an exon. The different choice of 5' or 3' splice site can be altered the length of exons in the alternative 5' and alternative 3' splice sites. The intron retention is the splicing failure to remove an intron. Finally, mutually exclusive type is the multiple cassette exons which are spliced by the different exon skipping in several choices of exons (Black 2003, Srebrow and Kornblihtt 2006). Such a complex patterns of alternative splicing produces numerous splicing variants including insertion or deletion of exons, frame shift, and premature termination codon (Sorek et al 2004). The choice of exons in alternative spicing is regulated by interaction between *cis*-acting elements and *trans*-acting factors.

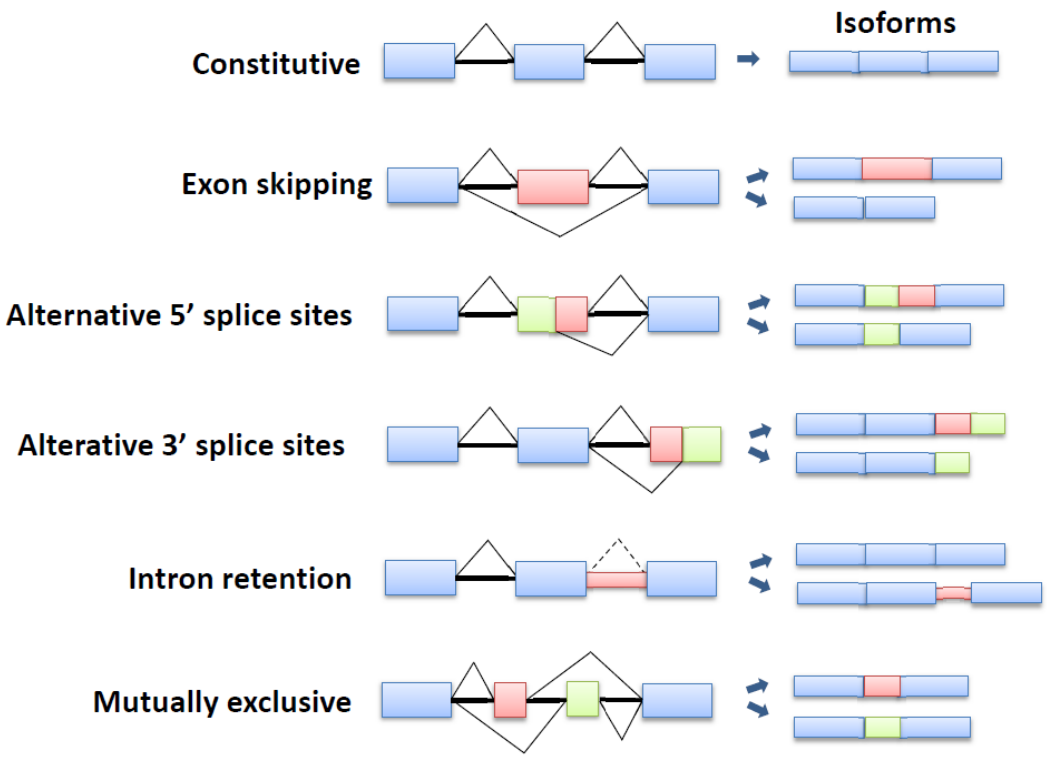

Figure 1-8. Different patterns of alternative splicing. Constitutive splicing produces mRNA including all exons and excluding all intron. Exon skipping or cassette exon is the common splicing pattern which includes or excludes exon. Alternative 5' or 3' splice sites patterns utilizes a different splice site instead of a normal splice site. Intron retention is a failure of intron that usually produces a frame shift or a premature stop codon. Mutually exclusive is a series of exon skipping by the selection of an exon from many exons (modified from Srebrow and Kornblihtt 2006).

1.10 *Cis*-acting elements in the alternative splicing

Regulation of alternative splicing is similar to the transcription regulation and involves a *cis*-acting element (enhancer and silencer sequence) and trans-acting factor (repressor and activator protein factor). Splicing regulatory elements is named by its physical location such as exonic splicing enhancer (ESE), intronic splicing enhancer (ISE), exonic splicing silencer (ESS), and intronic splicing silencer (ISS) (Figure 1-9). These elements interact with trans-acting factors to activate or suppress the neighboring splicing sites (Black 2003, Matlin et al 2005). In case of splicing enhancer, purine rich enhancer and A/C-rich enhancer (GAR repeat) have been identified (Coulter et al 1997). In terms of ESE and ISE's function, splicing enhancer support weak alternative splice sites and assist spliceosome and trans-acting factors. Compared to splicing enhancer, splicing silencer is not well understood but ESS and ISS probably recruits splicing repressor protein to suppress spliceosome assembly (Srebrow and Kornblihtt 2006).

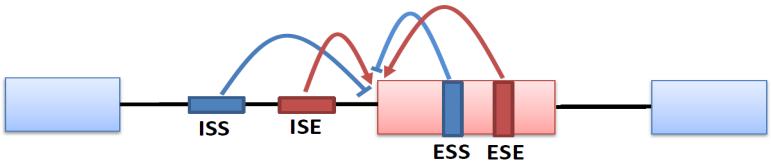

Figure 1-9. *Cis*-acting elements in the alternative splicing. Exonic splicing enhancer (ESE) and intronic splicing enhancer (ISE) activates spliceosome assembly, indicated by red arrows. Exonic splicing silencer (ESS) and intronic splicing silencer (ISS) suppresses spliceosome, represented by blue lines (modified from Srebrow and Kornblihtt 2006).

1.11 *Trans*-acting factors in the alternative splicing

There are 4 classes of trans-acting factors that regulate the alternative splicing. The primary factors are U1, U2, U4, U5, and U6 snRNPs which are spliceosomal components, interacting with 5' splice site, 3' splice site, polypyrimidine tract, and a branch point (Figure 1-7).

The second group of factors is a serine/arginine-rich protein (SR protein) family, typically binding to the ESE (Figure 1-10). SR proteins are essential pre-mRNA splicing factors containing highly conserved the RNA-recognition motif (RRM motif) in N-terminal and arginine/serine domain (RS domain) in C-terminal (Graveley 2000). The RRM motif plays a role in RNA binding and protein-protein interactions with U1snRNP and U2AF, suggesting that SR protein regulates splice site selection (Boukis et al 2004, Wu and Maniatis 1993).

The third group is heterogeneous nuclear ribonucleoprotein (hnRNPs). The hnRNPs are the most abundant proteins in the nucleus. hnRNP have diverse functions including assisting trafficking of mRNA, associating with other splicing factors, mRNA stability, and polyadenylation of mRNA (Krecic and Swanson 1999). hnRNP is contains RRM motif for binding to mRNA and glycine-rich domain for protein-protein interaction (Gorlach et al 1993, Martinez-Contreras et al 2007). In case of hnRNP A1 act as a negative regulator which inhibit U1 snRNP in alternative splicing (Figure 1-10) (Eperon et al 2000).

Final group of trans-acting factor for alternative splicing is the RNA binding motif (RBM) family and the CUGBP and ETR-like factors (CELF) family of protein, which contains RNA binding domain to activate or suppress numerous target mRNA (Han and

Cooper 2005). Other factors including SR protein kinases, phosphatases, and methylases can be engaged in the trans-acting factors.

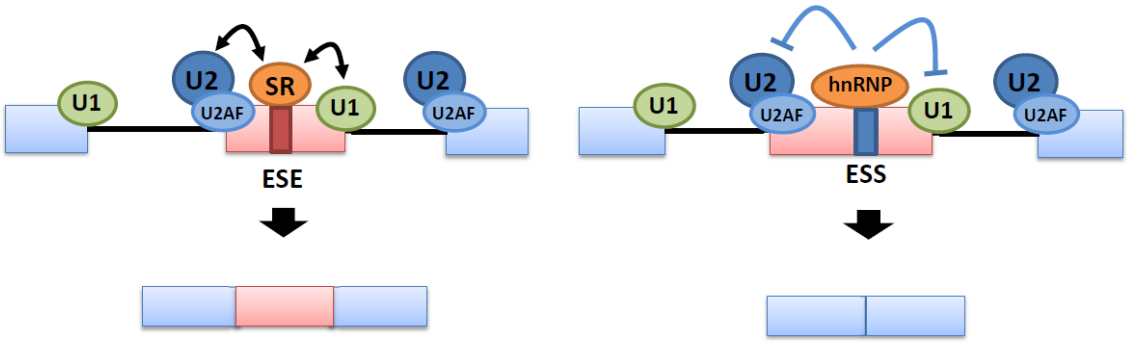

Figure 1-10. Trans-acting elements in the alternative splicing. The basic function of SR proteins binds to the exonic splicing enhancer (ESE) and promotes exon definition including recruitment of U2AF and U1 by protein-protein interaction (Left panel). hnRNPs bind to the exonic splicing silencer (ESS), and then inhibit exon definition by blocking the interaction between U2AF and the 3' splice site (Right panel). In addition, hnRNPs directly compete with SR protein to repress the binding of U1 and U2 to the exon (modified from House and Lynch 2008).

1.12 alternative splicing in cancer

A number of studies provide the importance of an alternative splicing in human diseases. Mutation of splicing regulatory elements on an exon or intron affects the splicing event in cancer (Grosso et al 2008, Venables et al 2009, Wang and Cooper 2007).

For examples, two point mutations in exon 7 (V211 G>A) and intron 13 (A>T) of the BRCA2 gene splicing sites produced a truncated protein in its C-terminal region (Pensabene et al 2009). A splice site mutation of c-Met gene exon14 produces the juxtamembrane domain deletion in gastric cancer cell line Hs746T (Asaoka et al 2010). Mutations in cis-acting elements, such as ESE and ESS, can also disturb alternative splicing. For example, a point mutation of the APC tumor suppressor gene in an ESE site (the exon 14 1918C>G) perturbed the binding sites of SRp55, hnRNP A1 or ASF/SF2 splicing factor. As a result, exon 14 of APC was completely skipped, leading to a frame shift and a premature termination codon (Goncalves et al 2009).

Furthermore, mutations can generate new splicing regulatory elements. For example, a point mutation in intron 1 of the KLF6 gene generates a new binding site for the SR protein SRp40 (Narla et al 2005). KLF6 acts as a tumor suppressor gene and is somatically inactivated in prostate cancer (Narla et al 2001). The new SRp40 binding site produces KLF6 splicing variants which antagonize the biological functions of wild type KLF6.

There are several examples of alternative splicing isoforms affect to human cancer including p53, fibronectin, fibroblast growth factor (FGF) receptor, vascular endothelial growth factor (VEGF) receptor, CD44, and murine double minute (MDM) 2 (Hofstetter et al 2010, Rennel et al 2009, Srebrow and Kornblihtt 2006). Furthermore, some of the

alternatively spliced isoforms can accumulate in tumors and are sufficient to promote cell transformation and metastasis formation in cell culture (Klingbeil et al 2009, Singh et al 2004).

Alteration of SR factors and hnRNP proteins frequently occur in tumors and are accompanied by alterations in the relative abundance of alternative splicing products, a typical signature of cancer cells (David et al 2010, Ghigna et al 2008). Cancer cell lines are frequently characterized by a high level of alternative splicing events that are not conserved between human and mouse and are not found in physiological tissues (Kan et al 2005).

The alternative splicing of CD44, a trans-membrane glycoprotein involved in cell to cell and cell to matrix interactions, is a good example of how alternative splicing in cancer can be modulated by trans-acting factors (Naor et al 2002). The production of different CD44 isoforms correlates with changes in the abundance of SR proteins (Huang et al 2007). Several CD44 isoforms have been identified diverse inclusions in proximal extra-cellular domain of CD44 exons (v1-v10). Wild type CD44 has no alternative exons and predominantly expressed in normal tissues. However, CD44 isoforms, containing variant exons v5, v6 and v7, are over-expressed in various tumors and have been implicated in tumor cell invasion and metastasis (Afify et al 2008, Naor et al 2002).

The functional importance of trans-acting factors in cancer cells reported that splicing factor SF2/ASF acts as a proto-oncogene and several target genes of SF2/ASF are essential for the oncogenic activity of S6K1, which is a novel oncogenic isoform of the mTOR substrate (Karni et al 2007). In addition, up- or down-regulation of RNA binding

proteins in cancer is associated with alternative splicing of oncogenes or tumor-suppressor genes (Grosso et al 2008).

1.13 RBM5 expression in cancer

RNA binding proteins are involved in diverse RNA metabolism including RNA splicing, transport, translation, and stability. So far, 15 different RNA binding motif (RBM) proteins have been identified which contains one to four copies RBM or RNA recognition motif (RRM), serine/arginine-rich domain or arginine/glycine-rich domain (Sutherland et al 2005). RBM family proteins are involved in alternative splicing, apoptosis, and cancer. For example, RBM3 is down-regulated in polyglutamine tract-induced apoptosis (Kita et al 2002) and RBM4 is involved in the splice site selection (Lai et al 2003). RBM7 may play a role in cell and tissue specific RNA processing (Guo et al 2003). In addition, RBM8 is involved in spliced mRNA transport and the process of nonsense-mediated decay (NMD) of mRNA containing premature stop codon (Kataoka et al 2000, Maquat and Carmichael 2001). Although biological functions of many RBM proteins are still unknown, the biological function of RBM5 is relatively well studied. RBM5, also known as LUCA-15 or H37, is an RNA binding protein and deletion of this gene is found in human lung cancer (Oh et al 2002). Recently, it was found that RBM5 regulates alternative splicing between exon-8 and -10 of caspase 2, leading to production of more pro-apoptotic splicing forms (Fushimi et al 2008). Moreover, alternative splicing of Fas pre-mRNA produces a membrane-bound receptor that promotes apoptosis and a soluble isoform that prevents apoptosis; Fas splicing is also regulated by RBM5, which is

a partner of U2AF, an important splicing factor for early splice-site recognition (Bonnal et al 2008, Kotlajich and Hertel 2008). In contrast, it has been reported that expression of RBM5 is increased in breast and ovarian cancers, correlated with HER-2/neu proto-oncogene (Rintala-Maki et al 2007). Thus, the exact biological activity of RBM5 in tumorigenesis needs to be elucidated, and investigation of the alternative splicing mechanism of the individual gene is required.

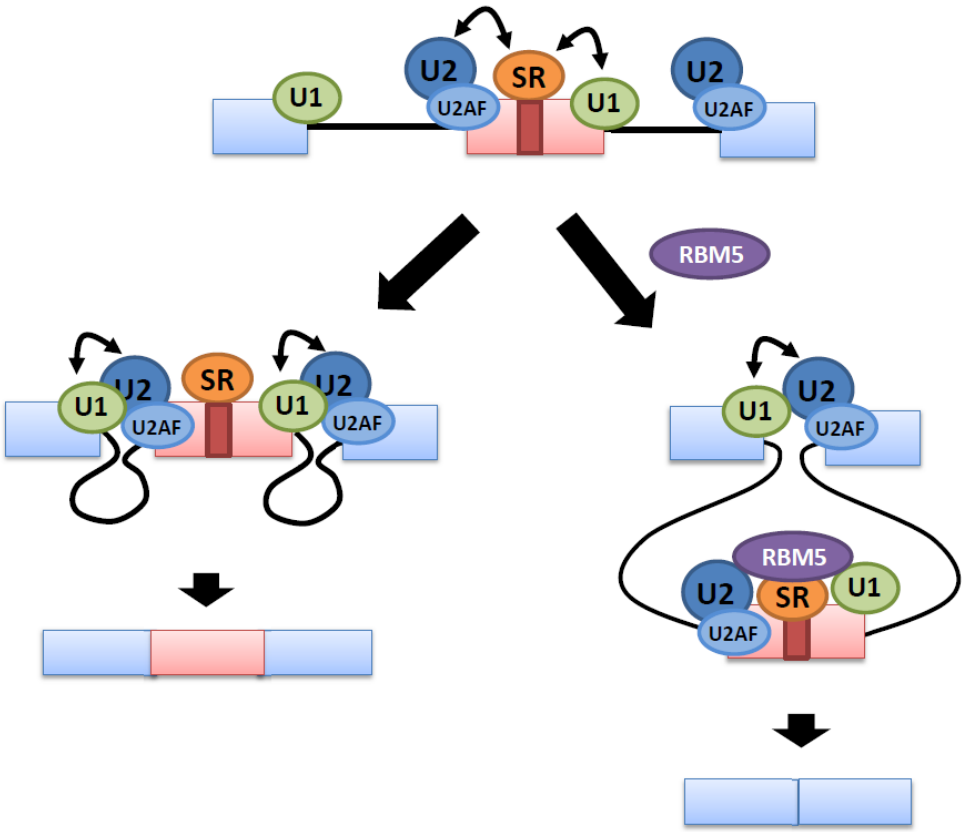

Figure 1-11. Regulation of RBM5-mediated alternative splicing. U1 snRNP and U2AF recognize both 5'and 3' splice sites. Consequently, SR protein stabilizes cross-exon interactions, resulting in constitutive alternative splicing (Left). In the presence of RBM5, the interaction between SR protein and spliceosomal components is masked, thus promoting exon skipping by allowing cross-intron interaction (Right). Figure is modified from Kotlajich and Hertel 2008.

1.14 Summary

As discussed in chapter 1, alternative splicing may be a major contributor to tumorigenesis. The objective of this study was to identify the mechanism of KLF4 alternative splicing in colorectal cancer. Based on preliminary data, we hypothesized that alternative splicing of KLF4 is regulated by RBM5 in colorectal cancer. Detailed approaches are described in chapter 3. Briefly, we identified novel alternative splicing variants of KLF4 in several human cancer cell lines. Sequence analysis revealed that the smallest KLF4 splicing variant is produced by exon skipping. In addition, we developed the KLF4 minigene system to investigate the exon skipping mechanism of KLF4 and found that regulation of alternative splicing of KLF4 is mediated by RBM5. Furthermore, sulindac sulfide (SS) treatment altered KLF4 alternative splicing by the dephosphorylation of RBM5 in colorectal cancer cells, providing a novel mechanism of sulindac sulfide effect in anti-cancer activity. Overall, our studies offer a novel mechanism to regulate tumorigenesis by alternative splicing of KLF4.

CHAPTER 2

Materials and

Experimental Procedures

2.1 Cell lines

Human colorectal cancer cell lines (HCT-116, SW480, HT-29 and LoVo) and other cell lines were purchased from American Type Culture Collection (Manassas, VA). SqCC/Y1, which is head and neck cancer cell line, was obtained from Dr. Dong M. Shin (Emory University, Atlanta, GA). HCT-116 and HT-29 were maintained in McCOY's 5A media. LoVo, A549 (lung cancer cell), PC-3 (prostate cancer cell) was incubated in Ham's F12 media. SW480, H292 (lung cancer cell), and SqCC/Y1 was grown in RPMI 1640 media. MCF-7 (breast cancer cell) was incubated in Dulbecco's Modified Eagle media (DMEM). T98G (brain cancer cell) was maintained in EMEM media. CMT-93 (mouse coloreactal cancer cell) was grown in DMEM media. 3T3-L1 (mouse embryonic fibroblast, pre-adipocyte) was cultured in DMEM with 10 % bovine calf serum. SCC7 (mouse squamous cell) was maintained in MEM media. MEF (mouse embryo fibroblast) was grown in DMEM media. All media, except DMEM for 3T3-L1 cell line, were supplemented with 10% (v/v) fetal bovine serum (FBS), and 100U/ml penicillin and 100 μg/ml streptomycin at 37 °C with 5% CO_2.

2.2 Molecular biology Kits

Human colon tumor and normal adjacent tissue pairs total RNA was purchased from Ambion (Austin, TX). RNA isolation kit was purchased from Omega Bio-Tek (Norcross, GA). iScript cDNA synthesis kit was purchased from Bio-Rad (Hercules, CA). pGEM-T easy cloning vector, Green Taq polymerase, and *in vitro* translation kit were obtained from Promega (Madison, WI). Agarose gel extraction kit and miniprep kit were purchased Zymo Research (Orange, CA). The QuikChange II site-directed mutagenesis

kit was obtained from Stratagene (La Jolla, CA). pCR 2.1 TA TOPO cloning, pcDNA3.1 V5-tagged TOPO and pcDNA3.1/CT-GFP-tagged TOPO expression vector were obtained from Invitrogen (Carlsbad, CA). Enhanced Chemiluminescence (ECL) kit for western blot and Chemiluminescence nucleic acid detection kit for EMSA was obtained from Pierce (Rockford, IL). Nuclear extract kit for EMSA was purchased from Active Motif (Carlsbad, CA). Primers, listed in table 3, were purchased from Operon (Pittsburgh, PA).

2.3 Antibodies

KLF4, RBM5, and β-Actin primary antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). Serine/threonine phospho antibody was purchased from BD biosciences (San Jose, CA). Antibody for V5 was purchased from Invitrogen (Carlsbad, CA). Secondary antibodies for western blot and tetramethyl rhodamine iso-thiocyanaste (TRITC) labeled secondary antibody for immunofluorescence were purchased from Southern Biotech (Birmingham, AL).

2.4 Chemicals

Sulindac sulfide (SS) was purchased from Calbiochem (San Diego, CA). 5-[[6-[(2-fluorophenyl)-methoxy]-2-napthalenyl]methyl]-2,4-thiazolidinedione (MCC-555) was obtained as a gift from Mitsubishi Pharma Corporation (Tokyo, Japan). Rosiglitazone (RGZ) and Tolfenamic acid (TA) were purchased from Cayman Chemical Co (Ann Arbor, MI). 3,3'-diindolylmethane (DIM), and (-)-epigallocatchin-3-gallate (EGCG) were purchased from Sigma (St. Louis, MO). 4',6-diamidine-2'-phenylindole

dihydrochloride (DAPI) for nuclear staining was obtained from Roche Applied Science (Indianapolis, IN). All other chemicals were purchased from Thermo Fisher Scientific (Pittsburgh, PA).

2.5 Total RNA isolation

Total RNA was isolated using the RNA isolation kit according to the manufacturer's instructions. In case of reagent treatment such as MCC-555, rosiglitazone, TA, sulindac sulfide, DIM, and EGCG, the cells were grown until 60-70% confluent in 6-well plates. The cells were incubated with serum free media overnight, and then reagents were treated in serum free media for 4h or 24h. Next, the cells were washed with phosphate buffered saline (PBS) and harvested with 350µl of cell lysis buffer. After total RNA isolation, RNA concentration was measure by spectrophotometer at 260/280nm.

2.6 Reverse Transcription PCR (RT-PCR)

Total RNA (1µg) was reverse-transcribed with the iScript cDNA kit according to the manufacturer's instructions. Briefly, 1 µg of total RNA was mixed with 5X iScript reaction mixture including reverse-transcriptase, dNTP, oligo-dT, and random primers in PCR tubes. The cDNA synthesis reaction was as followed; 25°C for 5min, 42°C for 30min, 85°C for 5min, and 4°C for holding. After cDNA synthesis, 1/20 of cDNA was used for the template and PCR was performed using 2X Green Taq polymerase with primers (see Table 3). In order to perform the PCR reaction for KLF4, the PCR condition was as followed; 94°C for 30 s, 55°C for 30 s, and 72°C for 2 min for 25 cycles for KLF4

or 20 cycles for GAPDH, which is internal control. After RT-PCR reaction, samples were loaded into 1.0% agarose gel with 0.5X TAE buffer (Tris-HCl/acetate/EDTA buffer) and ethidium bromide, which is a DNA intercalating agent, to visualize DNA under UV. The agarose gel was separated at 100V for 20min and the intensity of PCR bands were quantified by using Scion Image software (Frederick, MD). Each sample was normalized to the intensity value of *GAPDH*.

2.7 KLF4 cDNA constructs

$KLF4_L$, $KLF4_M$, and $KLF4_S$ cDNAs were amplified by RT-PCR using human total RNA from Ambion. Each KLF4 PCR bands were isolated using agarose gel extraction kit. 100 ng of cDNAs were subcloned into pcDNA 2.1 TOPO cloning vector and then transformed into DH5α competent bacterial cells on Luria-Bertani (LB) plate containing 100μg/ml ampicillin. After overnight incubation at 37°C, several single colonies were inoculated into LB broth with 100μg/ml ampicillin and incubated in the bacterial shaker at 37°C for overnight. Plasmid DNA was purified using Miniprep kit from Zymo Research and cDNA insert was verified by DNA sequencing. *KLF4* cDNAs were digested with *EcoR* I and *Hind* III and then subcloned into V5-tagged-pcDNA3.1 expression vector (Invitrogen). DNA sequence was confirmed by sequencing.

2.8 RBM5 cDNA construct

RBM5 cDNAs were were amplified by RT-PCR using human total RNA from Ambion. Next, the cDNA was subcloned into pGEM-T easy cloning vector. cDNA was verified

sequencing. *RBM5* cDNA was digested with *BamH* I and *Xba* I and then, ligated with pcDNA 3.1 expression vector. Subcloning procedures occurred as stated.

2.9 KLF4 minigene construct

The KLF4 minigene splicing cassette was generated by amplifying a 1.6 kb genomic DNA fragment from the *KLF4* gene (exon-2 to exon-4) with *Pfu* polymerase using human genomic DNA (Promega, WI). KLF4 genomic fragment was subcloned into pcDNA3.1 CT-GFP-TOPO expression vector digested with *Kpn* I and *Xba* I sites and then verified by sequencing.

2.10 Site-directed mutagenesis

Point mutations of *KLF4* minigene and *RBM5* were produced by the site-directed mutagenesis kit (Stratagene, Santa Clara, CA). Briefly, the reaction mixture was as follow; 10 ng of template DNA, 125 ng of forward and reverse primers, dNTP, and *Pfu* polymerase. The reaction mixture was incubated under the following conditions; 94°C for 1 min (initial step), 94°C for 1min, 58°C for 1min, and 68°C for 9 min for 25 cycles for *KLF4* and 68°C for 10 min for 25 cycles for *RBM5*. Next, 10 U of *Dpn* I restriction enzyme was added into the PCR mixture and incubated at 37°C for 1 h. Subsequently, 1/5 sample was used for transformation into XL-1 blue bacterial competent cells. Each mutagenesis was carefully confirmed by sequencing.

Table 3. Primer sequence for $KLF4_L$, $KLF4_S$, RBM5, and the KLF4 minigene construct.

Primer	Oligo sequence
KLF4 *EcoR* I F	5'-CGAATTCTATGGCTGTCAGCGACGCG-3'
KLF4 *Hind* III R	5'-CCCAAGCTTTTAAAAATGCCTCTTCATGTGTAAGGC-3'
$KLF4_S$ F	5'-ATGGCTGTCAGCGACGCGCTG-3'
$KLF4_S$ R	5'-TTAAAAATGCCTCTTCATGTGTAAGGC-3'
mKLF4 F	5'-ATGGCTGTCAGCGACGCTCTG-3'
mKLF4 R	5'-AAAGTGCCTCTTCATGTGTAAGGC-3'
GAPDHF	5'- GGGCTGCTTTTAACTCTGGT-3'
GAPDHR	5'-TGGCAGGTTTTTCTAGACGG-3'
KLF4 mini F	5'-TGGCTGTCAGCGACGCGCTGCTCC-3'
KLF4 mini R	5'-CTGTGTGGGTTCGCAGGTGTGCCTTG-3'
KLF4miniMF	5'- GTCTCTCCCGCCCTGTTCTTGCAGCGCTTGG-3'
KLF4miniMR	5'-CCAAGCGCTGCAAGAACAGGGCGGGAGAGAC-3'
RBM5 *BamH* I F	5'-CCGGATCCGGGACAATGGGTTCAGACAAAAGAGTG -3'
RBM5 *Xba* I R	5'-CCTCTAGACTCTCCATCTCAGTGAACCGGGC -3'
RBM5S69AF	5'-GAAAGAAGGAACGCTGACCGATCCGAAG-3'
RBM5S69AR	5'-CTTCGGATCGGTCAGCGTTCCTTCTTTC-3'

2.11 Transient transfection and luciferase reporter assays

Transient transfection was performed using PolyJet DNA *in vitro* transfection reagent from SignaGen (Ijamsville, MD), according to the manufacturer's instruction. Briefly, HCT-116 cells were plated in 12-well plates at 10^5 cells per well and were grown for 16 h. Plasmid mixture containing PAI-1, cyclin D1, or TOPFlash luciferase constructs, and pRL–null (Promega, WI) were transfected and the cells were fed fresh medium (McCoy's 5A medium with 10% FBS) overnight. After cell harvest using 1X luciferase lysis buffer, luciferase activity was determined and normalized to the *pRL-null* luciferase activity using a dual luciferase assay kit (Promega, WI).

2.12 siRNA experiment

RBM5 siRNA was obtained from Santa Cruz Biotechnology (Santa Cruz, CA). siRNA treatment was performed using TransIT-TKO transfection reagent from Mirus Bio (Madison, WI). According to manufacture protocol, HCT-116 cells were plated into 6 well plates at 10^6 per well and incubated for 16h. Next, the reaction mixture, containing 10 µl TKO reagents in 250 µl serum free McCoy 5A media, was incubated for 5 min at RT and then, 100 nM RBM5 siRNA was added with additional 20 min incubation. The siRNA mixture was directly added into HCT-116 cells and incubated 48h.

2.13 Establishment of stable cell line

HCT-116 cells were plated in 6-cm plates and transfected with KLF4 minigene or CONTROL vector (pcDNA3.1/V5-His-TOPO/LacZ) using LipofectAMINE (Invitrogen,

Carlsbad, CA), according to the manufacturer‘s protocol. After 24 h, the cells were then transferred to a 10 cm plate with G418 (800μg/ml) (Stratagene). Selection with G418 was carried out for 3 weeks and then RT-PCR was carried out to check the stable cell lines.

2.14 Western blot analysis

The cells were grown to 60-80% confluence in 6-cm plates followed by 24 h treatment in the presence of indicated compounds. The cells were then washed once with PBS and harvested using ice-cold RIPA buffer (PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease inhibitors (1 mM PMSF, 5 μg/ml aprotinin and 5 μg/ml Leupeptin) and phosphatase inhibitors (1 mM Na_3VO_4 and 1 mM NaF). Cell lysates were collected and kept on ice for 30 minutes followed by protein collection by quick centrifugation. The protein concentration was determined by BCA protein assay (Pierce, Rockford, IL) using bovine serum albumin (BSA) as the standard. Equal aliquots of 30 μg protein were boiled in 2X loading buffer (0.1 M Tris-Cl, pH 6.8, 4% SDS, 0.2% Bromophenyl blue, 20% glycerol) for 5 min, and then separated by SDS-PAGE (8-14% gels), followed by transfer to nitrocellulose membranes (Osmonics, Minnetonka MN) for 1 h. The blots were blocked for 1 h with 5% skim milk in TBS/Tween 0.05% (TBS-T), and probed with a specific primary antiserum in tris buffered saline (TBS) containing 0.05% Tween-20 (TBS-T) and 5% non-fat dry milk at 4°C overnight. After washing with TBS-T, the blots were treated with horseradish peroxidase-conjugated secondary antibody for 1 h and washed several times. Proteins were detected by the enhanced chemiluminescence system (Pierce).

2.15 Nickel pull-down analysis

HCT-116 cells were transiently transfected with either RBM5 expression vector (pcDNA3.1/V5-His-TOPO/RBM5) or control vector (pcDNA3.1/V5-His-TOPO/LacZ) using Lipofectamin2000 (Invitrogen) according to the manufacturer‘s protocol**.** The RBM5 proteins containing six histidine residues at their C-termini were purified using ProBond nickel-chelating resin (Invitrogen), under native condition buffer (50 mM NaH2PO4/pH 8.0, 500 mM NaCl, 1 mM PMSF, 1 μg/ml aprotinin, 1 μg/ml leupeptin), and subsequently incubated with 50 μL of ProBond Resin (50% Slurry in 20% ethanol) for 30 min at 4°C. Beads were then washed three times, boiled for 5 min, and loaded on SDS/PAGE for Western blot analysis.

2.16 Immunofluorescence

Immunofluorescence was performed as described previously (Lee et al 2008). Briefly, the cells were washed with PBS, fixed with 4% paraformaldehyde for 30 min, and washed three times with PBS for 5 min. The cells were incubated with anti-V5 primary antibody overnight at 4°C and then with goat anti -mouse TRITC conjugate for 1 h at room temperature in the dark. The cells were stained with 0.5 mg/ml of DAPI for 10 min to counterstain the nucleus. The expression of proteins was detected using a Nikon Eclipse E600 fluorescence microscope. The TIF images were captured using QCapture software version 2.66.4 with X400 magnification.

2.17 Electrophoretic mobility shift assay (EMSA)

$KLF4_L$ and $KLF4_S$ proteins were prepared following the manufacturer's protocols (Promega, WI) using the $KLF4_L$ and $KLF4_S$ pGEM-T easy construct. Synthetic oligonucleotides corresponding to the KLF4 binding site contained the following sequence: 5'-ATGCAGGAGAAAGAAGGGCGTAGTATCTACTAG-3' (Shields and Yang 1998). EMSA was performed using a Light Shift Chemiluminescent EMSA kit following the manufacturer's protocols (Pierce, IL). Briefly, biotin-labeled oligonucleotide (10 nM) was incubated with purified nuclear extract or *in vitro* translated KLF4 (5 μg), 1X binding buffer, 2.5% glycerol, 5 mM $MgCl_2$, 50 ng/μl Poly (dI·dC), and 0.05% NP-40 at room temperature for 20 min. For competition assay, nuclear extracts were preincubated with the unlabeled oligonucleotide (10X and 100X) for 10 min. DNA-protein complexes were resolved by 5% nondenaturing polyacrylamide gel and transferred to a nylon membrane, followed by chemiluminescent nucleic acid detection according to the manufacturer's protocols.

2.18 Human normal and adjacent tumor tissues

These samples were generously provided by Dr. Joo-Heon Yoon (Yonsei University, Korea). This study was approved by the Institutional Review Board of Yonsei University Health System, and all specimens were supported by the Colon Cancer Specimen Bank from the National Research Resource Bank Program of the Korea Science and Engineering Foundation in the Ministry of Science and Technology. For comparison, three histologically normal colons and three adjacent to colorectal adenocarcinomas were selected from the colon cancer patients who were never treated with any therapeutic purpose.

2.19 Statistical analysis

Statistical analysis was performed using the unpaired Student *t* test. *p<0.05, **p<0.01, ***p<0.001

CHAPTER 3

Results:

Alternative splicing of KLF4

in colorectal cancer

3.1 KLF4 splicing variants in human and mouse cancer cells

To investigate the expression of KLF4 mRNA in human and mouse cell lines, we performed RT-PCR with specific primers covering the coding region of KLF4 mRNA. The results revealed that at least three KLF4 splicing variants were produced in human colorectal and non-colorectal cancer cell lines (Figure 3-1A). Hereafter, we designate the KLF4 splicing variants as large KLF4 ($KLF4_L$), medium KLF4 ($KLF4_M$), and small KLF4 ($KLF4_S$). The highest amount of $KLF4_L$ mRNA expression was observed in LoVo cells, and the lowest amount of $KLF4_L$ was in A549 and PC-3 cells. However, all the cancer cell lines showed a similar amount of $KLF4_S$ mRNA expression. We also tested KLF4 splicing variants in mouse cell lines by RT-PCR (Figure 3-1B). RT-PCR analysis revealed that the mouse KLF4 gene generates two splicing variants: $mKLF4_L$ and $mKLF4_S$, indicating that alternative splicing of KLF4 is not only seen in humans but also in mice.

Since KLF4 protein expression was significantly decreased in gastric and colorectal cancer (Dang et al 2000a, Wei et al 2005, Zhao et al 2004), we determined whether KLF4 splicing is different in normal and adjacent tumor tissues. Therefore, we attempted to investigate $KLF4_L$ and $KLF4_S$ mRNA expression patterns by RT-PCR in human colon normal and tumor tissue. $KLF4_S$ transcripts were significantly detected in both normal and tumor samples, but the $KLF4_L$ transcript level was significantly reduced in tumors (Figure 3-2).

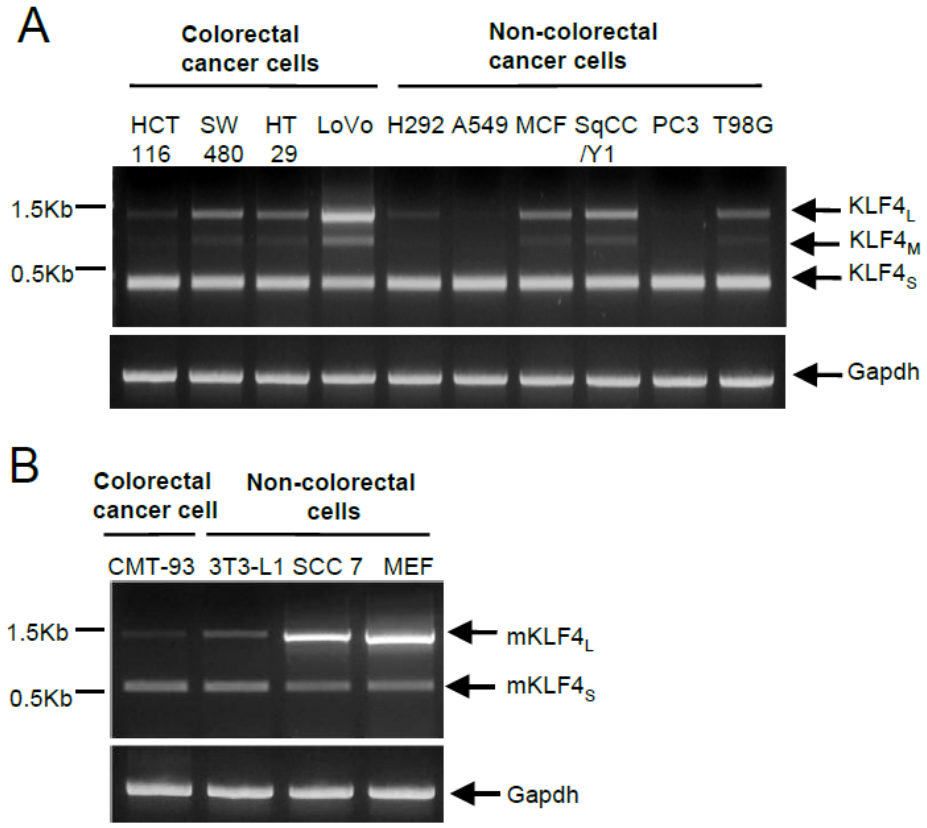

Figure 3-1. KLF4 splicing variants in different cell lines. (A) Three major human KLF4 splicing variants, KLF_L, KLF_M, and KLF_S, were detected by RT-PCR. RT-PCR analysis was performed with total RNA extracts from human cancer cell lines. Colorectal cancer cells are HCT-116, SW480, HT-29, and LoVo. Non-colorectal cancer cells are H292 (lung), A549 (lung), MCF7 (breast), SqCC/Y1 (head and neck), PC3 (prostate), and T98G (brain). (B) RT-PCR analysis of KLF4 splicing variants in mouse cell lines. CMT-93 is a mouse rectal cancer cell. Mouse non-colorectal cancer cells are 3T3-L1 (fibroblast), SCC7 (squamous cell carcinoma), and mouse embryonic fibroblast (MEF). Equal loading was confirmed by determining *Gapdh* transcript.

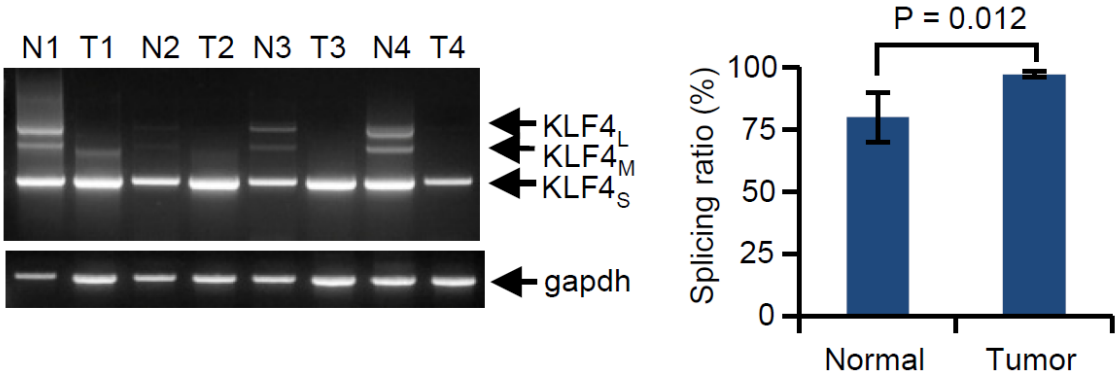

Figure 3-2. $KLF4_L$, $KLF4_M$, and $KLF4_S$ in human normal colon and tumor tissue samples. Human normal tissue (N1) and adjacent human tumor tissue RNAs (T1) were purchased from Ambion. N2-4 and T2-4 total RNAs were isolated from three colorectal cancer patients. KLF4 transcripts were analyzed by RT-PCR, and PCR products were confirmed by sequencing. Quantitative analysis from the RT-PCR is shown in the graph (*Right*). Error bars indicate standard deviation. The formula for the splicing ratio used in this study is as follows. Splicing ratio = $[KLF4_S/(KLF4_L+KLF4_S)]\times100$

3.2 $KLF4_S$ was produced by exon skipping

To obtain further information of KLF4 splicing variants in humans and mice, we cloned individual RT-PCR products and verified them by sequencing. Sequence analysis revealed that the human $KLF4_S$ transcript encodes a 154 amino acid protein by exon skipping (Figure 3-3). The skipped KLF4 exon-3 is the biggest exon, containing a transcription activation domain and an inhibitory domain (Wei et al 2006). Interestingly, human $KLF4_S$ contains the nuclear localization signal (NLS) and three C2-H2 zinc finger domains (Figure 3-4). In case of mouse KLF4, $mKLF4_S$ was produced by combination of 5' splice site and 3' splice site alternative splicing patterns. Also, $mKLF4_S$ is missing a signal peptide and one zinc finger domain (Figure 3-5). It is possible that $mKLF4_S$ could not translocate to the nucleus compared to human $KLF4_S$ (Figure 3-6).

Another KLF4 splicing variant $KLF4_M$ is produced by the alternative 3' splice site mode. However, $KLF4_M$ contains a premature termination codon that probably leads to nonsense-mediated decay (NMD) (Brogna and Wen 2009).

Sequence analysis suggests that human $KLF4_S$ might be able to translocate to the nucleus and bind to KLF4 target gene as an antagonist. Genomic DNA sequence analysis of the 5' and 3' splice site regions between human and mouse KLF4 revealed that the second intron 3' splice site may be important to exon skipping as a consequence of different $KLF4_S$ splicing patterns between two species (Figure 3-3).

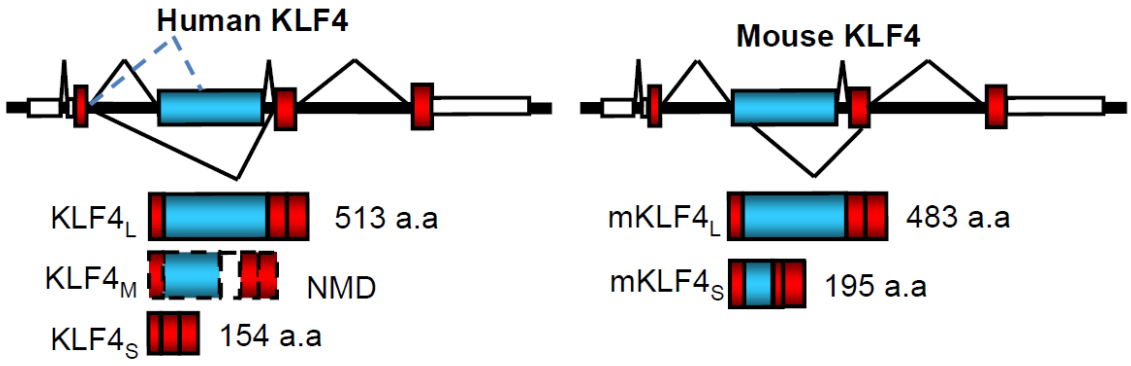

Figure 3-3. $KLF4_S$ is produced by exon skipping. Schematic representation of human and mouse KLF4 splicing variants. Human $KLF4_L$ and $KLF4_S$ encode 513 and 154 amino acids, respectively, by exon skipping. $KLF4_M$ contains a premature stop codon and could be degraded by the NMD pathway. Mouse $KLF4_S$ ($mKLF4_S$) are produced by a combination of the 5' and 3' alternative splicing sites. The open box indicates the 5' and 3' untranslated region (UTR). The blue box represents the 3^{rd} exon, skipped to generate $KLF4_S$.

Figure 3-4. Amino acid sequence alignment of human $KLF4_L$ and $KLF4_S$. Sequence alignment was done by using Clustral W2 (http://www.ebi.ac.uk/Tools/clustalw2). The number indicates amino acid length of KLF4. Nuclear localization signal (NLS) is indicated by the blue box, and the C2H2 zinc finger domains are indicated by red boxes.

Figure 3-5. Amino acid sequence alignment of mouse $KLF4_L$ and $KLF4_S$. Sequence alignment was done by using Clustral W2 (http://www.ebi.ac.uk/Tools/clustalw2). The number indicates amino acid length of KLF4. Nuclear localization signal (NLS) is indicated by the blue box, and the C2H2 zinc finger domains are indicated by red boxes.

Figure 3-6. Amino acid sequence alignment between human $KLF4_S$ and mouse $KLF4_S$. Sequence alignment was done by using Clustral W2 (http://www.ebi.ac.uk/Tools/clustalw2). The number indicates amino acid length of KLF4. The blue box indicates human $KLF4_S$ N-terminal region which encoded by exon 1 and 2. The red box represents human $KLF4_S$ C-terminal region. In mouse KLF4, red letters and blue letters indicate amino acids joined by combination of 5' splice site and 3' splice site alternative splicing.

3.3 Mutation of the 3' splice site in the second intron decreased the KLF4 splicing ratio

To investigate whether the 3' splicing site of the second intron affects human KLF4 exon skipping, we established the KLF4 minigene system (Figure 3-8A). The minigene system is a powerful tool to investigate the *cis*-acting element or *trans*-acting factor involving the alternative splicing event of the multiple-exon gene (Cooper 2005). The human KLF4 genomic fragment, from KLF4 exon-2 to exon-4 containing intron-2 and intron-3, was subcloned into pcDNA 3.1 CT-GFP vector. Alterative splicing was examined by RT-PCR. In addition, three nucleotides on the 3' splicing site of intron-2 in the human KLF4 gene were mutated to the mouse sequence (Figure 3-7). The wild type and mutated minigenes were individually transfected into HCT-116 cells and alternative splicing was determined by RT-PCR using vector-specific primers to discriminate endogenous KLF4 transcripts. The KLF4 minigene system effectively produced the alternative splicing variants $KLF4_L$ and $KLF4_S$, but not $KLF4_M$. Mutation of the 3'splice site in the second intron significantly suppressed the KLF4 splicing ratio, compared to wild type KLF4 in the minigene system (Figure 3-8B). This suggests that the 3' splicing site of the second intron contains an important *cis*-acting element for exon skipping of KLF4.

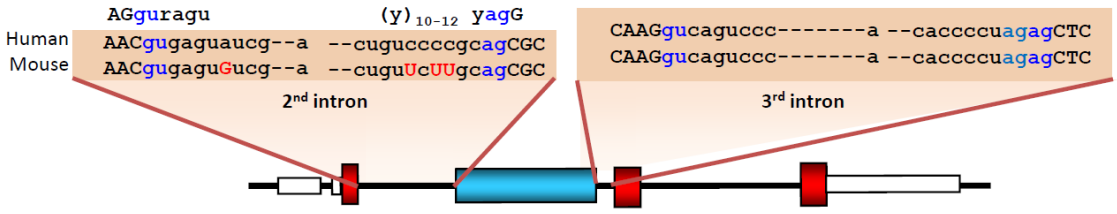

Figure 3-7. Sequence comparison of the splicing border regions between human and mouse KLF4 genes. Blue letters indicate the 5' and 3' splice site and "-a-" designates the branch site. Red letters represent different sequences between human and mouse KLF4.

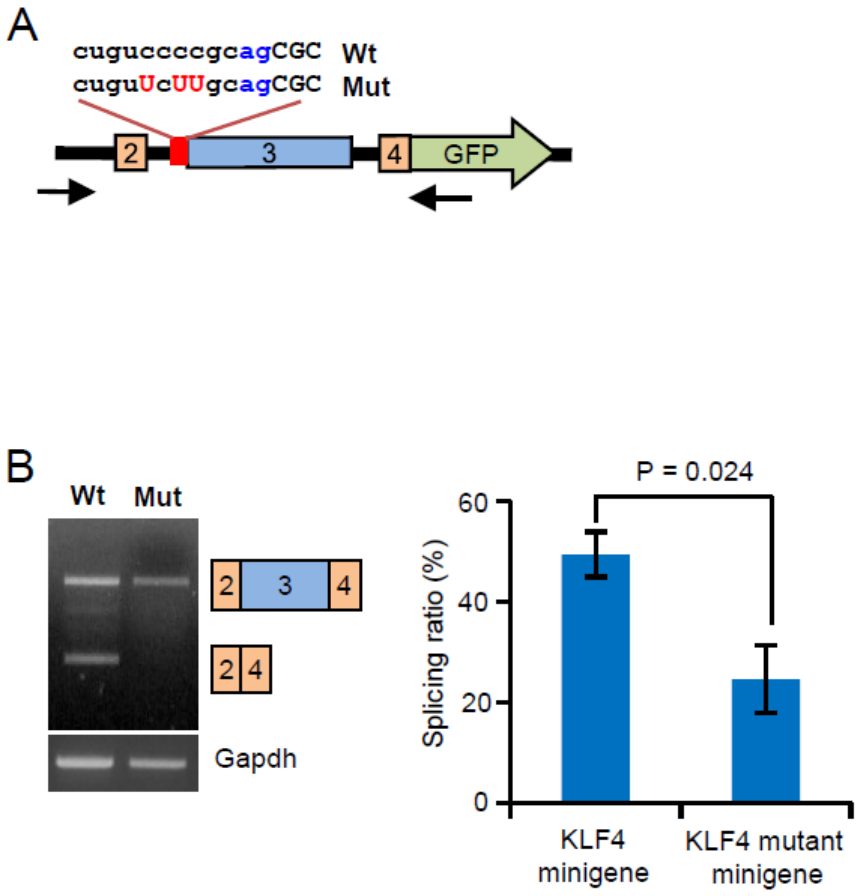

Figure 3-8. KLF4 3'splice site suppressed alternative splicing of the KLF4 minigene. (A) Schematic diagram of the KLF4 minigene, which contained the KLF4 exon 2-4 and intron 2-3. Three nucleotides on the 3' splice site within the second intron of human KLF4 were substituted to the mouse KLF4 3'splice site, indicated by red colors and a red line on the diagram. Arrows indicate primers for determining the splicing rate of the KLF4 minigene. (B) Mutation of the KLF4 3'splice site suppressed alternative splicing of the KLF4 minigene. The KLF4 minigene was transfected into HCT-116 cells and RT-PCR was performed. The graph represents mean values of KLF4 splicing ratios from three independent experiments. Error bars, SD.

3.4 Double 3' splice sites contribute to the in-frame transcripts of KLF4

One of the interesting features of KLF4 sequences is that the KLF4 carries a unique double 3' splice sites ("agag") on the intron-3 (Figure 3-9). The exon-3 inclusion on $KLF4_L$ utilizes the first "ag", whereas the exon inclusion on $KLF4_S$ utilizes the second "ag". This accurate process prevents the frame shift, resulting in identical C-terminal sequences in both $KLF4_L$ and $KLF4_S$ forms. For example, if $KLF4_L$ uses the second "ag" for connecting exon-3 and exon-4, this transcript produces 397 a.a rather than 513 a.a KLF4 protein by premature stop codon. On the contrary, if $KLF4_S$ utilizes the first "ag" for joining of exon-2 and exon-4, the transcript will be 109 a.a rather than 154 a.a protein. This is not only seen in human colorectal cancer cell lines, but also seen in $KLF4_L$ and $KLF4_S$ splicing variants from human tissue samples (Figure 3-2). To demonstrate the majority of 3' splice sites of the KLF4 exon-3 exclusion, RT-PCR products were carefully investigated by the direct PCR sequencing and minor usage of 3' splice sites can be abolished by using this method (Figure 3-9B).

A double 3' splicing site (agag) in FAS gene (CD95) from autoimmune lymphoproliferative syndrome (ALPS) patients has been reported (Roesler et al 2005). However, the double 3' splice site in KLF4 gene is a natural sequence rather than a point mutation as seen in FAS gene. Therefore, further studies are necessary to investigate the regulation mechanism of the KLF4 double 3' splice sites.

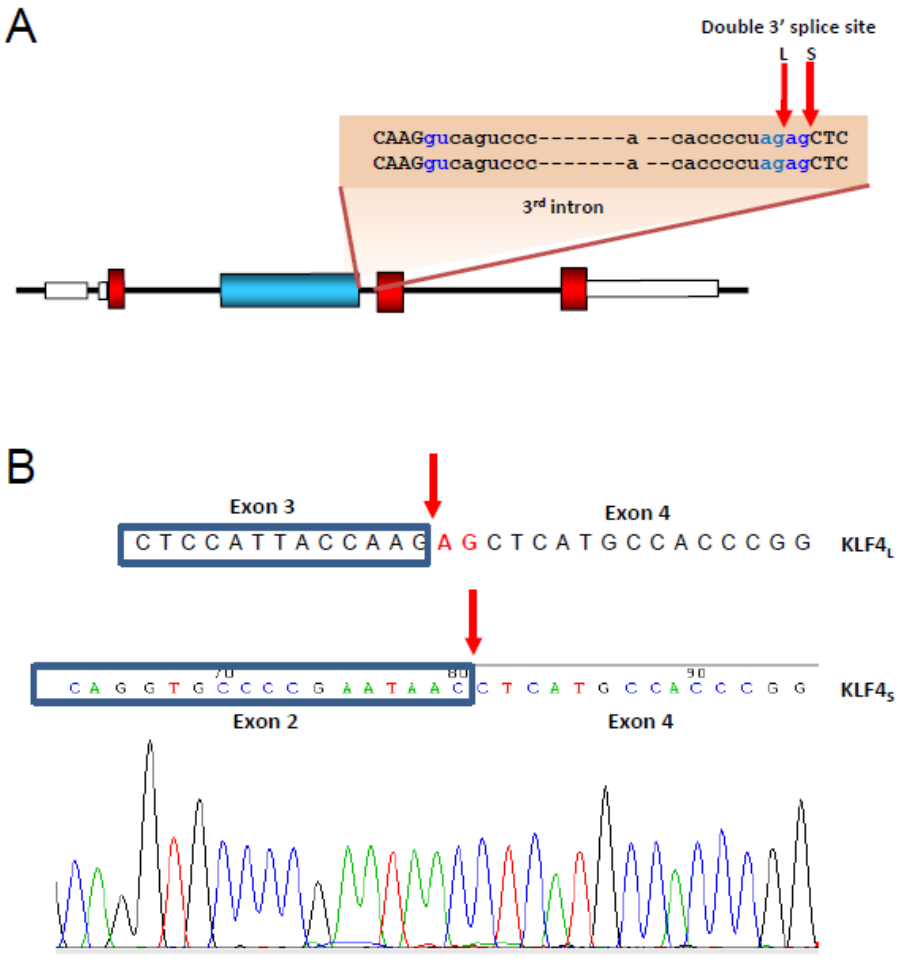

Figure 3-9. Double 3' splice sites on intron-3 of KLF4. (A) The double 3' splice sites are indicated by red arrows; L is the 3'splice site for $KLF4_L$ and S is for $KLF4_S$. (B) Sequencing data for the exon joining region on the $KLF4_S$. $KLF4_L$ contains "ag", indicating red letters and red arrow shows the border between exon-3 and exon-4. Alternative spicing of $KLF4_S$ produces without "ag" on border between exon-2 and exon-4, indicated by red arrow.

3.5 RBM5 altered the KLF4 splicing ratio

Recent studies showed that RBM5 regulates alternative splicing of FAS and caspase 2 mRNA (Bonnal et al 2008, Fushimi et al 2008). Therefore, we questioned whether KLF4 alternative splicing is also controlled by RBM5. RBM5 expression vector and KLF4 minigene were co-transfected into HCT-116 cells and then alternative splicing was examined by RT-PCR. Overexpression of RBM5 significantly increased KLF4 alternative splicing ratio, suggesting that RBM5 controls KLF4 splicing (Figure 3-10A). On the other hand, the down-regulation of RBM5 by siRNA treatment decreased the KLF4 splicing ratio (Figure 3-10B). Based on these results, RBM5 could be a major mediator to control KLF4 alternative splicing.

3.6 Sulindac sulfide (SS) restores $KLF4_L$ in HCT-116 cells via RBM5 dephosphorylation

Previously, we reported that various nonsteroidal anti-inflammatory drugs (NSAIDs), PPARγ ligands, and dietary compounds have anti-tumorigenic effects in colorectal cancer cells (Sukhthankar et al 2008, Yamaguchi et al 2006). Therefore, we tested whether treatment with these compounds affects alternative splicing of KLF4. Among those, SS treatment (10 μM) significantly increased endogenous $KLF4_L$ and $KLF4_M$ in HCT-116 cells (Figure 3-11A). Using the KLF4 minigene system, we found that SS and MCC-555 suppress the splicing ratio of KLF4 in both HCT-116 and LoVo cells (Figure 3-11B), respectively; however, RBM5 overexpression increased splicing of the KLF4 minigene in

both cell lines. These results indicate a pivotal role of RBM5 in KLF4 alternative splicing in both cells. However, SS and MCC-555 may not be enough to fully increase $KLF4_L$ forms in the presence of RBM5.

Previous report also suggested that RBM5 is a phospho-protein; however, there is no evidence connecting RBM5 phosphorylation and splicing activation onto its target pre-mRNA (Ekim et al 2003). To test SS effects on the phosphorylation of RBM5, we performed nickel pull-down followed by western blot using the phospho-serine/threonine antibody. After SS treatment, total serine/threonine phosphorylation of RBM5 was decreased in a dose-dependent manner, whereas expression of total RBM5 was not changed (Figure 3-12A). Additionally, a point mutation of RBM5 at serine 69 (S69A) showed significant inhibition of KLF4 minigene splicing (Figure 3-12B). Therefore, our finding suggests that phosphorylation of RBM5 plays a role in KLF4 exon skipping and SS facilitates dephosphorylation of RBM5, leading to diminished exon skipping. This could provide a new mechanism of how SS treatment affects anti-tumorigenesis in colorectal cancer.

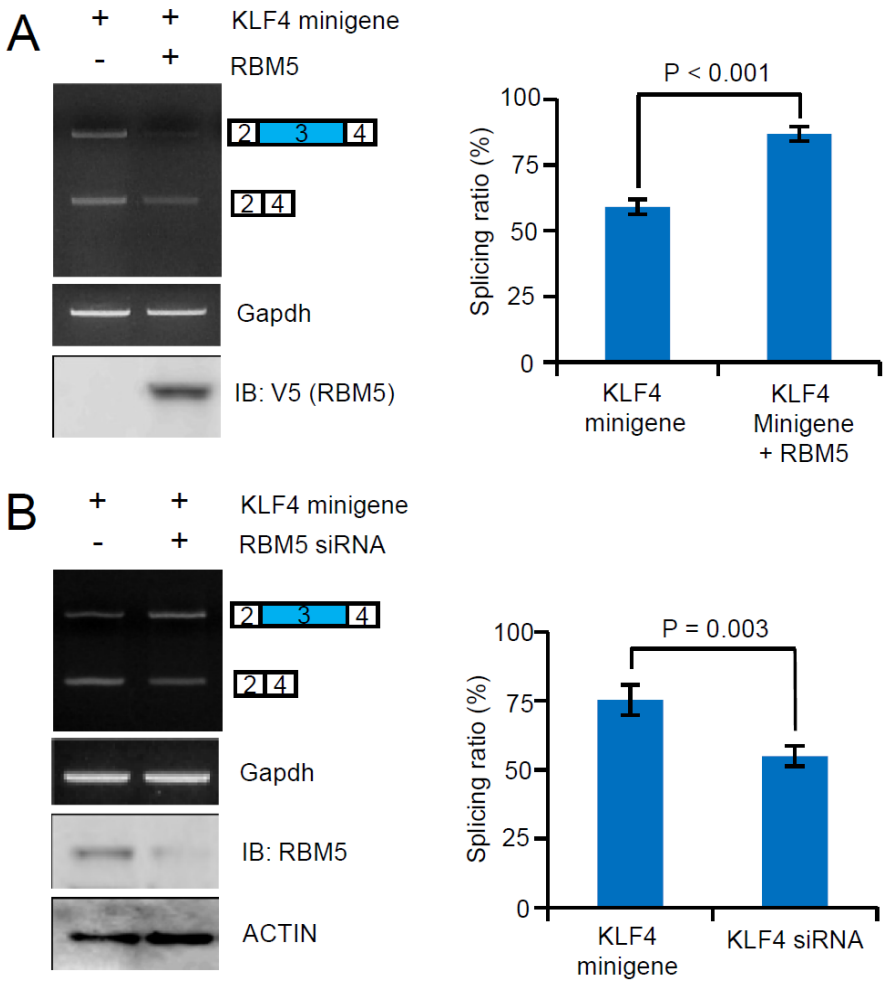

Figure 3-10. Overexpression or knockdown of RBM5 alters the KLF4 splicing ratio. (A) Overexpression of RBM5 increases alternative splicing of KLF4. KLF4 minigene and V5-tagged RBM5 construct were transfected into HCT-116, and then the KLF4 splicing ratio was examined by RT-PCR analysis. Overexpression of RBM5 was confirmed by Western blot analysis using the anti-V5 antibody. (B) Knockdown of RBM5 by siRNA decreases the KLF4 splicing ratio. RBM5 knockdown was confirmed by Western blot with RBM5 antibody. The bar graphs on the right represent mean value and the error bar indicates SD from three independent experiments.

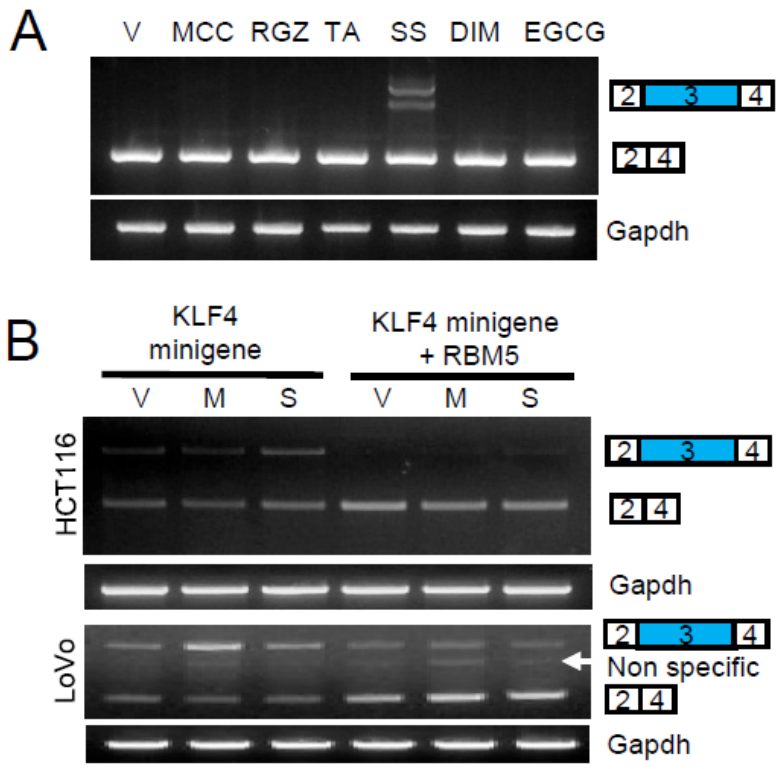

Figure 3-11. SS suppresses alternative splicing of KLF4. (A) HCT-116 cells were starved in serum free-media overnight and then incubated for 4 h with PPAR-γ ligands, NSAIDs, and dietary compounds; RNA was isolated for RT-PCR analysis. 10 µM SS treatment significantly restored endogenous $KLF4_L$ transcripts (V, vehicle; MCC, 10µM MCC-555; RGZ, 10µM Rosiglitazon; TA, 30µM Tolfenamic acid; SS, 10µM Sulindac sulfide; DIM, 10µM 3,3'-diindolylmethane; EGCG, 10µM epigallocatechin gallate). *Gapdh* is the loading control. (B) RBM5 overexpression enhanced KLF4 alternative splicing under the KLF4 minigene system. The KLF4 minigene was transiently transfected for 24 h into HCT-116 cells or LoVo cells and then the cells were treated with vehicle (V), 5 µM MCC-555 (M), and 10 µM SS (S) for 4 h.

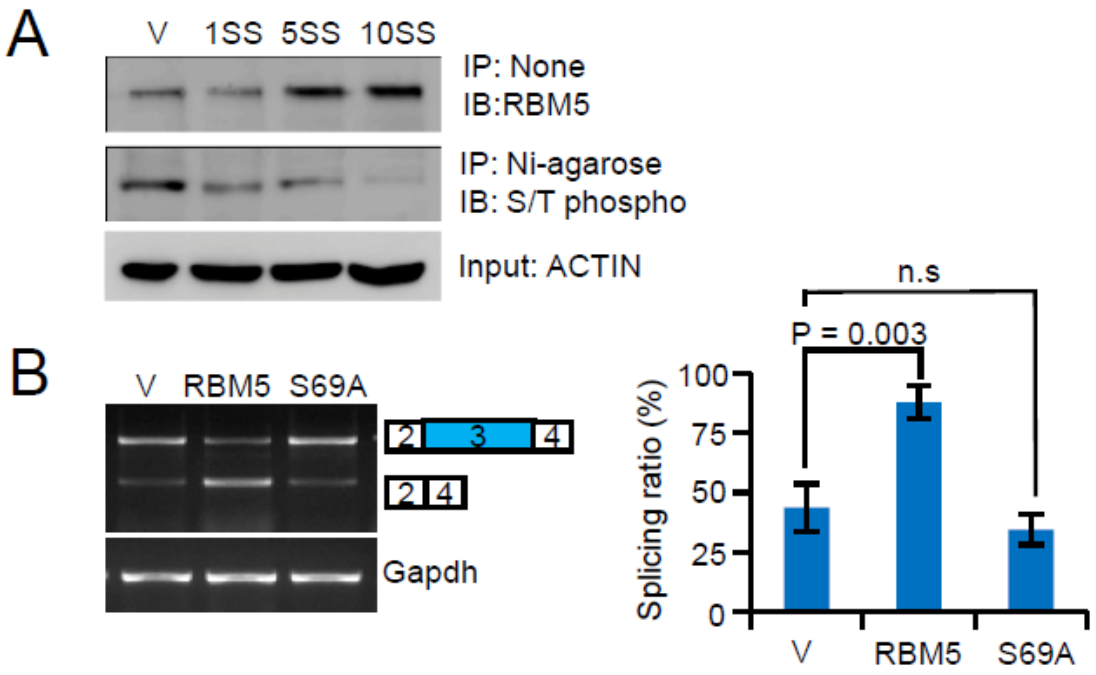

Figure 3-12. Dephosphorylation of RBM5 altered KLF4 splicing. (A) Phosphorylation of RBM5 was inhibited by sulindac sulfide treatment. His-tag RBM5 was pulled down by Ni-agarose bead and western blot was performed with phospho-serine/threonine antibody. Actin was the loading control. (B) RBM5 mutant (S69A) could not facilitate the KLF4 splicing. The wild type RBM5 and mutant RBM5 construct were cotranfected with the KLF4 minigene into HCT-116 and the splicing ratio measured by RT-PCR. The bar graphs on the right represent mean value, and the error bar indicates standard deviation, from three independent experiments.

3.7 Biological roles of $KLF4_S$

Significant amounts of $KLF4_S$ mRNA and sequence analysis results suggest that $KLF4_S$ could have particular biological activities in the cell. To elucidate biological function of $KLF4_S$ in human colorectal cancer cells, we overexpressed $KLF4_S$ into HCT-116 cells. The V5-tagged $KLF4_S$ was detected by Western blot analysis. The extra bands of overexpression of $KLF4_L$ could be degraded forms as previously observed (Shie et al 2000). In addition, endogenous $KLF4_L$ expression was examined by using the anti-KLF4 antibody, and all the transfected cells expressed the expected protein as assessed by Western blot (Figure 3-13).

Based on our sequence analysis, we determined $KLF4_S$ could be located in the nucleus since $KLF4_S$ contained a nuclear localized signal (Figure 3-4). To identify the subcellular localization of $KLF4_S$, we performed immunofluorescence using the anti-V5 antibody in HCT-116. The lacZ protein, as a control, was predominantly expressed in the cytoplasm, and $KLF4_L$ was highly expressed in the nucleus. However, $KLF4_S$ was detected both in the cytoplasm and nucleus (Figure 3-14A). To verify our immunofluorescence data, we isolated nuclear and cytoplasm fractions from overexpressed KLF4 cell lysates and then performed the Western blot with the V5-antibody. We found that lacZ was predominantly expressed in the cytoplasm, whereas $KLF4_L$ was predominantly expressed in the nucleus, consistent with the immunofluoroscence results; however, $KLF4_S$ was expressed in both cytoplasm and nucleus (Figure 3-14B).

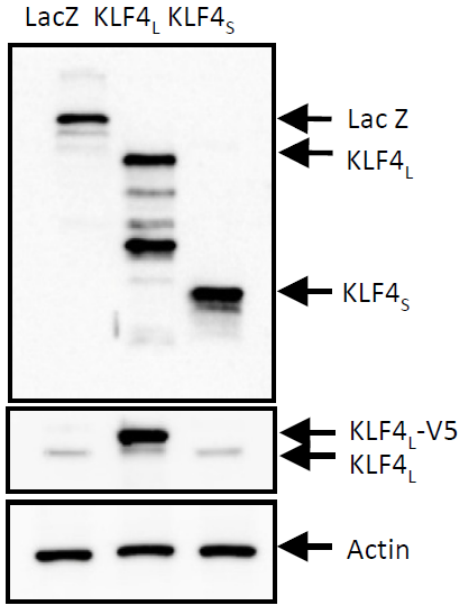

Figure 3-13. Western blot analysis of human $KLF4_L$ and $KLF4_S$ transiently transfected into HCT-116. After 24h Transfection, the cells were harvested with RIPA buffer. LacZ was used for the control experiment. Overexpressed proteins were detected by anti-V5 antibody (top). Endogenous KLF4 expression was detected by anti-KLF4 antibody (middle). Actin is the loading control.

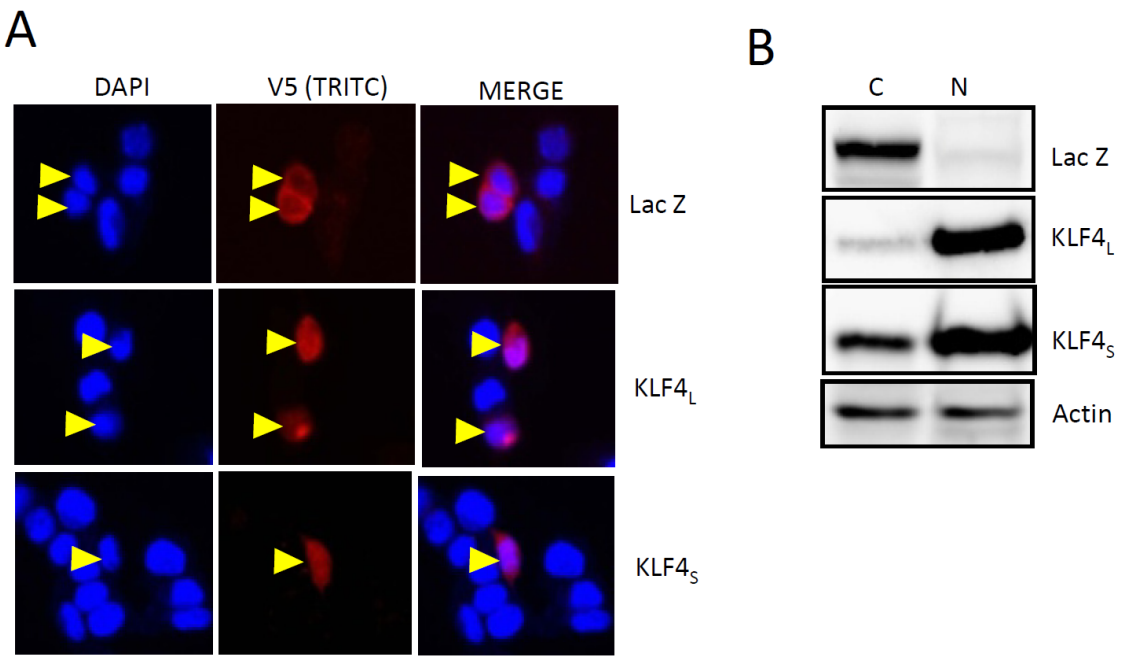

Figure 3-14. $KLF4_S$ has cellular localization and DNA binding activity similar to $KLF4_L$. (A) Localization of V5-tagged overexpressed $KLF4_L$ and $KLF4_S$ in HCT-116 cells was visualized by fluorescence microscope. The nucleus was stained by DAPI and TRITC for LacZ, $KLF4_L$, and $KLF4_S$. (B) Localization of $KLF4_L$ and $KLF4_S$ was confirmed by Western blot analysis using cytosol fraction (C) and nucleus fraction (N) from HCT-116. Actin is the loading control.

Finally, DNA binding activity of $KLF4_S$ was investigated by the electrophoretic mobility shift assay (EMSA) with biotin-labeled oligonucleotides for the KLF4 binding element (Shields and Yang 1998). Both $KLF4_L$ and $KLF4_S$ were obtained from the *in vitro* translation method and an apparent DNA-protein complex was observed (Figure 3-15). In addition, the V5-tagged KLF4/DNA complex was recognized on the same membrane by Western blot. These results suggest that $KLF4_S$ is localized in the nucleus and binds to the KLF4 binding element.

Cyclin D1 is a well known KLF4 target gene (Shie et al 2000). To elucidate $KLF4_S$ transcription activation or repression activity, a -163/+130 cyclin D1 promoter construct containing a single KLF4 binding site (Castro-Rivera et al 2001) was cotransfected with either $KLF4_L$, $KLF4_S$, or LacZ into HCT-116. As previously reported, overexpression of $KLF4_L$ suppressed luciferase reporter gene activity (Shie et al 2000). However, $KLF4_S$ did not repress *cyclin D1* promoter activity, compared to $KLF4_L$ (Figure 3-16). The *Plasminogen activator inhibitor-1* (*PAI-1)* promoter containing four KLF4 binding sites was also used to test KLF4s activity (Chen et al 1998). Cotransfection of $KLF4_L$ and the -884/+71 PAI-1 promoter construct repressed reporter gene transcription. However, KLF_S slightly inhibited reporter gene transcription compared to the control. These results suggested that $KLF4_S$ is an inactive form of $KLF4_L$.

A recent study showed that KLF4 interacts with the C-terminal of β-catenin, and inhibition of β-catenin signaling is required in both the N-terminal and C-terminal of the KLF4 domain (Wang et al 1996). $KLF4_S$ contains sufficient domains to bind with β-catenin, but not to inhibit β-catenin activity. Therefore, we tested TOPFlash reporter construct activity together with KLF_L and KLF_S (Figure 3-17A). As expected, $KLF4_L$

inhibited TOPFlash activity; however, KLF_S rather enhanced β-catenin signaling in HCT-116 cells. To identify whether $KLF4_S$ physically interacts with β-catenin, $KLF4_L$ and $KLF4_S$ was overexpressed in HCT-116 followed by Ni-agarose pull-down experiment (Figure 3-17B). Western blot results using β-catenin antibody showed that both $KLF4_L$ and $KLF4_S$ strongly bind to β-catenin. Therefore, $KLF4_S$ could modulate activity of $KLF4_L$ in Wnt/β-catenin signaling pathway in colorectal cancer.

Together with other results, $KLF4_S$ forms may inactivate $KLF4_L$'s tumor suppressor activity in colorectal tumorigenesis. This is evidence that alternative splicing may provide an important mechanism for the regulation of tumorigenesis in the human colorectal tumorigenesis.

3.8 Summary

Based on our results, we propose a model of the KLF4 alternative splicing mechanism (Figure 3-18), in which $KLF4_S$ is produced by exon skipping and antagonize $KLF4_L$ activities, and subsequently enhances tumorigenesis. Phosphorylation of RBM5 also enhances KLF4 alternative splicing toward $KLF4_S$; however, SS treatments repress KLF4 splicing processes by dephosphorylation of RBM5. Therefore, KLF4 splicing variants can be used as a potential diagnostic and prognostic marker for human colorectal cancer.

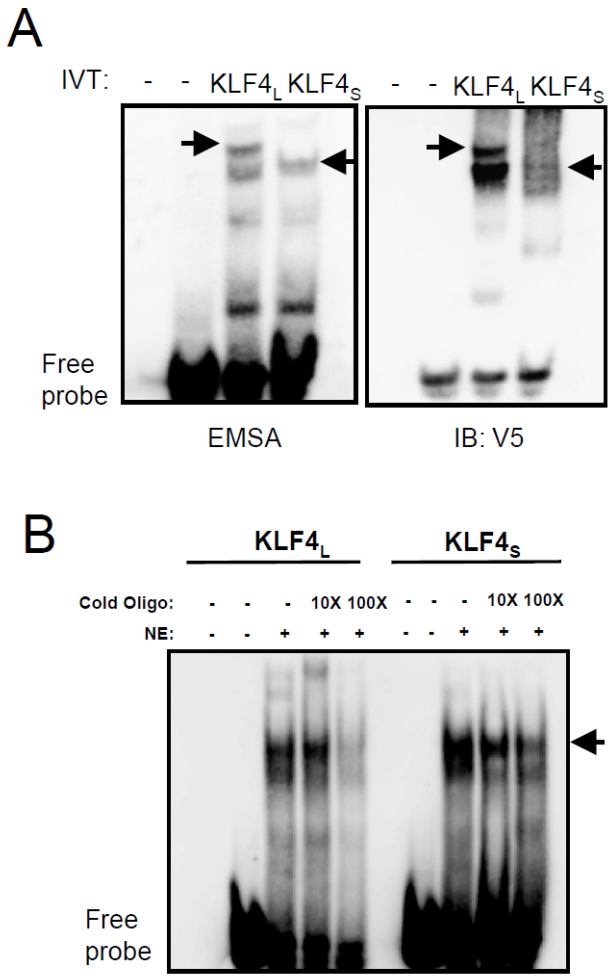

Figure 3-15. Electrophoretic mobility shift assay (EMSA) analysis of KLF_L and $KLF4_S$. (A) $KLF4_S$ is able to bind to the same $KLF4_L$ DNA binding site. EMSA analysis was performed using *in vitro* translated protein extracts. Western blot analysis evaluated recombinant KLF4 protein using V5 antibody to detect the DNA-protein complex followed by treatment of the stripping buffer (Thermo Scientific) on the same membrane. (B) EMSA was performed using nuclear extracts to confirm the DNA binding activity of $KLF4_L$ and $KLF4_S$. The competition of the DNA binding was obtained using a 10, and 100 time excess of the unlabelled oligonucleotide.

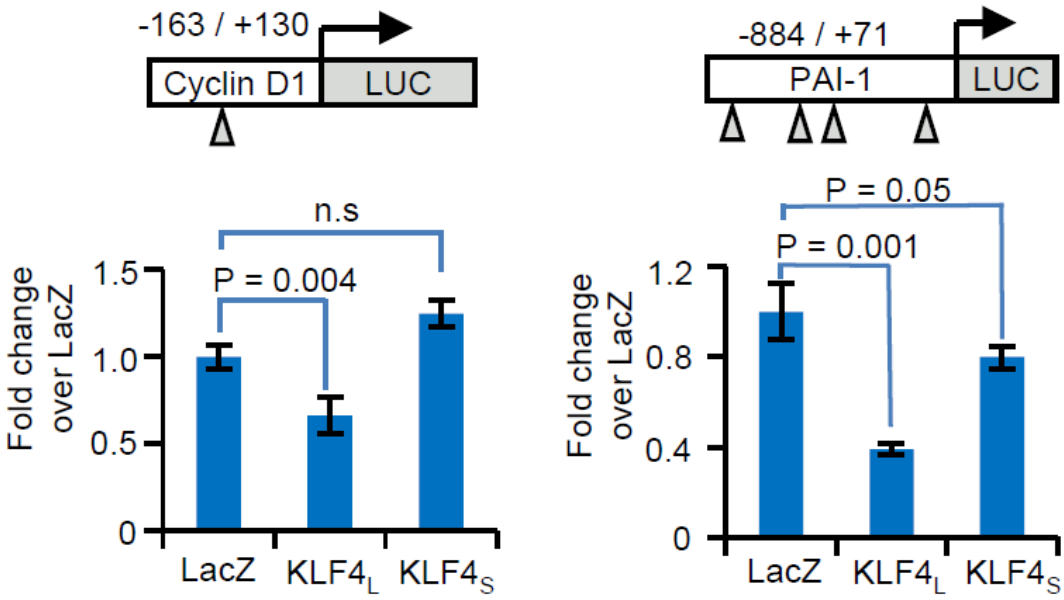

Figure 3-16. $KLF4_S$ antagonizes $KLF4_L$ activity. $KLF4_S$ suppresses $KLF4_L$ activity on the *cyclin D1* and *PAI-1* promoter. The upper diagram indicates *cyclin D1* and *PAI-1* luciferase reporter construct containing the KLF4 binding site, indicated by an arrowhead. $KLF4_L$ and $KLF4_S$ were co-transfected with *cyclin D1* or *PAI-1* construct into HCT-116 cells. Luciferase activity was measured as a ratio of firefly luciferase signal/renilla luciferase signal and is shown as mean ± SD of fold change over lacZ.

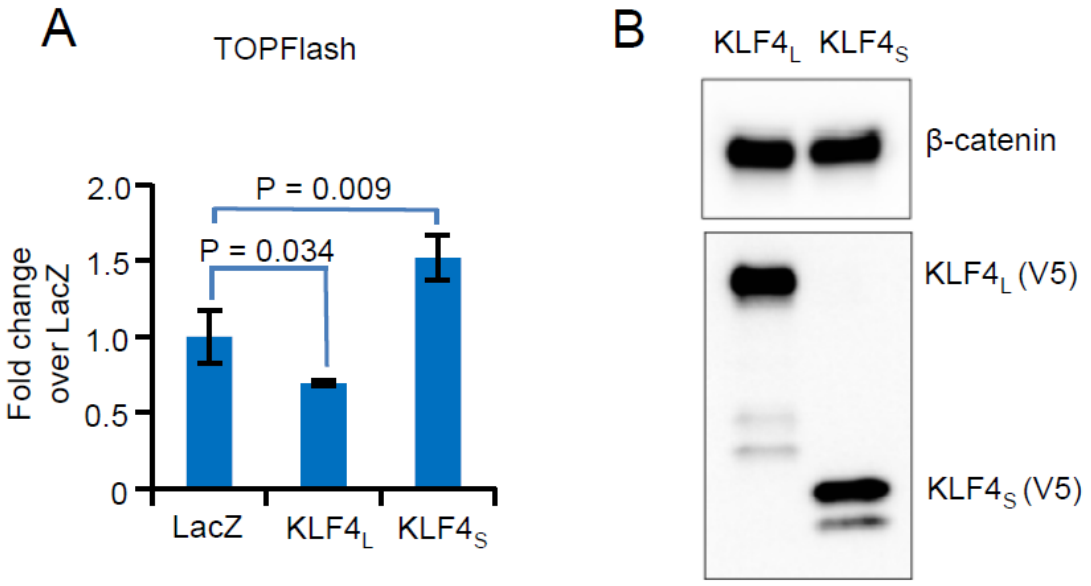

Figure 3-17. $KLF4_S$ interacts with β-catenin. (A) $KLF4_S$ is not able to inhibit TOPFlash reporter activity. The TOPFlash luciferase reporter construct was co-transfected with LacZ, $KLF4_L$, or $KLF4_S$ into HCT-116 cells. The represented results were independently performed three times. Error bars indicates SD. (B) Interaction between $KLF4_S$ and β-catenin was confirmed by Ni-pull down experiment. Both $KLF4_L$ and $KLF4_S$ were strongly interacted with β-catenin.

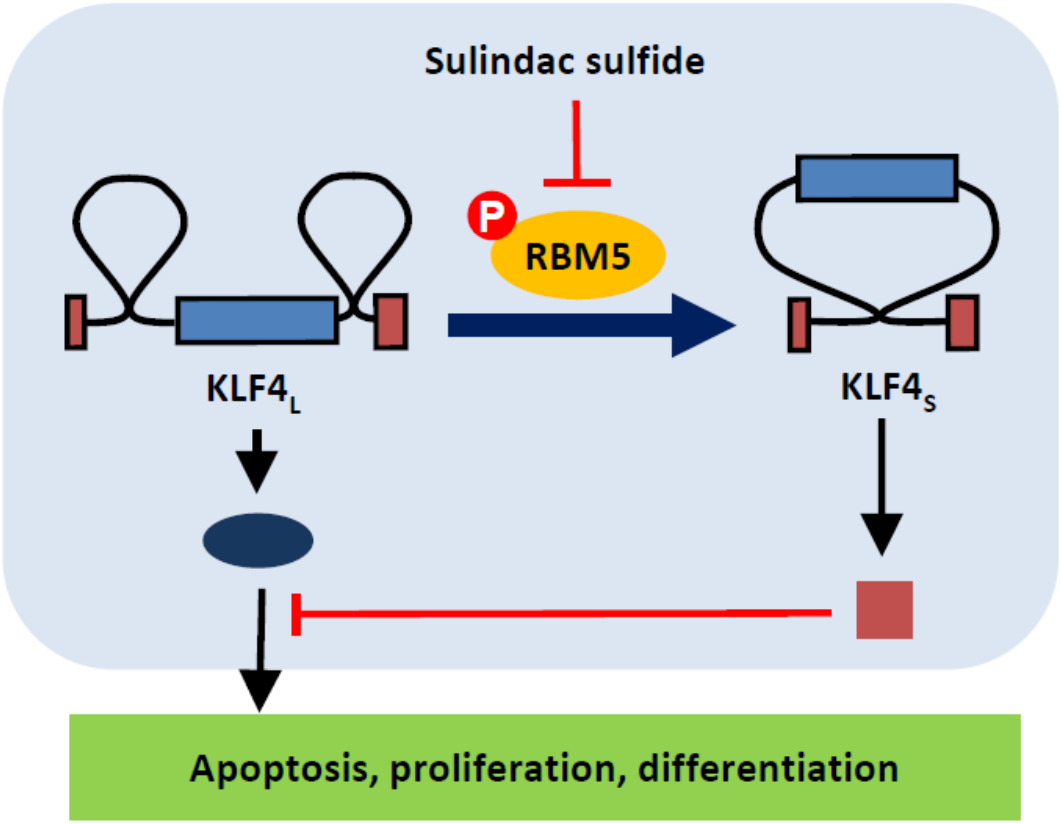

Figure 3-18. Model of alternative splicing of KLF4 and regulation by RBM5. In case of normal cells, $KLF4_L$ is associated with cell apoptosis, proliferation, and differentiation. The phosphorylated RBM5 facilitates KLF4 alternative splicing and generates $KLF4_S$, which inhibit $KLF4_L$ functions. In the presence of SS, the RBM5-mediated alternative splicing of KLF4 is inhibited and $KLF4_L$ mRNA production is restored in human colorectal cancer cells.

CHAPTER 4

Discussion

4.1 Alternative splicing occurs in KLF4 transcription.

Alternative splicing is an important post-transcriptional regulation mechanism to control normal and cancer cell processes including apoptosis and tumorigenesis. It has been reported that KLF4 protein level is significantly decreased in adenomas from FAP patients (Dang et al 2000a, Ton-That et al 1997) and, KLF4 mRNA level is decreased in adenomas and carcinomas, compared to normal tissue (Shie et al 2000). To our knowledge, this is the first report that alternative splicing (exon skipping) is observed during tumorigenesis in the *KLF4* gene. Three major variants of KLF4 are found in cancer cells: $KLF4_L$, $KLF4_M$ and $KLF4_S$ (Figure 3-1). In addition, $KLF4_S$ is generated by exon skipping and we could detect $KLF4_S$ mRNA in most cancer and normal cells. This obervation was seen in normal colon tissues when RT-PCR was performed using colon tissues (Figure 3-2). However, the level of $KLF4_L$ mRNA was barely detected in both cancer cell lines and human colon cancer tissue samples. Therefore, alterative splicing is a major contributor for regulating $KLF4_L$ mRNA level in tumorigenesis.

4.2 $KLF4_S$ antagonizes with $KLF4_L$ biological activity.

To examine the biological function of $KLF4_S$, we overexpressed $KLF4_S$ in human colorectal cancer cells. Interestingly, $KLF4_S$ can translocate to the nucleus and bind to the KLF4 binding site on KLF4 target genes including *cyclin D1* and *PAI-1*, and inhibit β-catenin signaling (Figure 3-16 and 3-17). Particularly, Wnt/β-catenin signaling is crucial for the maintaining and organizing of the human intestinal epithelium (Pinto et al 2003).

Recently, KLF4 directly interacts with β-catenin as well as T cell factor 4 (TCF4) to inhibit β-catenin and p300/CBP complex (Evans et al 2010). In Figure 3-17B, $KLF4_S$ directly interacts with β-catenin, suggesting $KLF4_S$ may play a pivotal role in tumorigenesis. Since KLF4 is a transcription factor and controls many genes involved in cell proliferation and differentiation, localization of $KLF4_S$ in the nuclei implies inhibition of KLF4 transcriptional activity not only in tumorigenesis, but also in differentiation. Indeed, KLF4 has been known to be involved in stem cell differentiation (Guo et al 2009). Further investigation is required to determine the biological significance of $KLF4_S$ in cell differentiation.

4.3 KLF4 splicing profile can be a useful diagnostic tool in human colorectal cancer.

$KLF4_L$ and $KLF4_S$ splicing variants were examined in human normal and tumor tissues by RT-PCR (Figure 3-2). It is interesting that all the normal tissues, but not adjacent tumor tissues, produced $KLF4_L$ and $KLF4_M$. This is consistent with our hypothesis that the 3' splicing site in the second intron plays an important role in the KLF4 splicing mechanism. Based on our sequence analysis, human and mouse KLF4 showed different 3' splice sites on the second intron, generating different splicing variants (Figure 3-8). It is possible that a weak 3' splice recognition site affects alternative splicing of mKLF4, because the full or partial exon-3 is found in both $mKLF4_L$ and $mKLF4_S$. Therefore, we developed the KLF4 minigene system and mutated the 3' splice site of intron-2. As expected, the substitution of the 3' splice recognition site from human to mouse KLF4

sequence suppressed the KLF4 splicing ratio (Figure 3-7). In addition, 3' splice site on the inton-3 contains the unique double "agag" sequence which showed the exon skipping dependent manner (Figure 3-9). Based on sequence analysis from human colon normal and tumor tissue sample, $KLF4_L$ utilizes the first "ag" and $KLF4_S$ uses the second "ag", which produces the same peptide sequences beyond Exon 3. This makes it difficult to generate a specific antibody for $KLF4_S$ peptide. However, if we generated a specific antibody based on this study, taken together with the identification of mutation at the splicing site, the KLF4 splicing profiles between normal and tumor samples could be used for diagnostic and/or prognostic markers for colon cancer in the future.

4.4 RBM5 binding site on the KLF4 mRNA

The identification of *cis*-acting elements (e.g. exonic splicing enhancer element and intronic splicing enhancer element) and *trans*-acting elements within pre-mRNA of KLF4 may give us insights into the relationship between the alternative splicing and tumorigenesis. To identify RBM5 binding sites on the intron-2 of KLF4, we performed several different experiments for RNA-protein interactions including RNA immunoprecipitation assay, RNA gel shift assay, and biotin labeled-RNA pull-down assay. Unfortunately, we could not find the specific RBM5 binding site on intron-2. It is thought that RBM5 may not directly bind to a specific sequence or other *cis*-acting elements located in a distant region may be required to provide optimal binding of RBM5 to the RNA sequence. To screen a potential RBM5 binding site on the KLF4 exon-3 which is excluded by the exon skipping, we analyzed sequence using the RESCUE-ESE

database (http://genes.mit.edu/burgelab/rescue-ese/). There are several potential ESE sequences to investigate for identification of RBM5 biding site using RNA gel shift assay and RNA-Chip assay. Since there is no conserved RBM5 binding sequence being identified, we cannot exclude the possibility that RBM5 may affect splicing without direct binding to the RNA sequences.

4.5 Therapeutic aspects of KLF4 splicing

Many pharmacological approaches have proposed the use of conventional drug therapy; oligonucleotide-mediated and RNA-based therapies could be valuable clinical applications for the alternative splicing-dependent human disease (Garcia-Blanco et al 2004). We screened dietary compounds and the following conventional drugs to examine whether they affect alternative splicing of KLF4: PPARγ ligands and NSAIDs. Endogenous $KLF4_L$ levels were especially increased by SS, a nonselective cyclooxygenase inhibitor: SS facilitates the dephophorylation of RBM5, which leads to inactive RBM5. Although the detailed mechanisms need to be elucidated, this is the first report suggesting that SS, a known colorectal cancer inhibitor, affects alternative splicing through the alteration of the spliceosome complex. However, a direct kinase for phosphorylation of RBM5 is still unknown. There are some reports showed that SR proteins are phosphorylated to have an effect on RNA-protein interaction and alternative splicing (Duncan et al 1997, Xiao and Manley 1997). So far, several kinases have been reported to phosphorylate SR proteins, including SR protein kinase (SRPK) family and

Cdc2-like kinase (Clk/Sty) family (Cao et al 1997, Jiang et al 2009, Mathew et al 2008, Ngo et al 2008, Zhong et al 2009). Possibly, SS is potential inhibitor of these kinases.

In contrast, RBM5 phosphorylation could be dephosphorylated by protein phosphatase 1 (PP1). Protein serine/threonine phosphatases or phosphoprotein phosphatases are Mg^{2+} dependent enzymes. Especially, PP1 is a major phosphatase in this family and very conserved protein in eukaryotes (Ceulemans and Bollen 2004). PP1 is associated with regulation of diverse cellular processes including cell signal transduction, stress response, protein synthesis, cell cycle arrest, and apoptosis (Ceulemans and Bollen 2004, Kwiek et al 2006). Moreover, PP1 is involved in alternative splicing by dephosphorylation of SR protein (Massiello and Chalfant 2006). Although there is no evidence that RBM5 is a target protein of PP1, SS may affect phosphatase activity of PP1 to regulate dephosphorylation of RBM5.

More than 100 different proteins are involved in spliceosome complex recognition of the exon-intron boundary, and dozens of serine-arginine factors enhance or repress the alternative splicing process (Ghigna et al 2008). However, effects of conventional drugs in the spliceosome complex and splicing process are not known.

4.6 Future directions for alternative splicing in cancer

Importance of alternative spicing in cancer research is emphasized by many studies (Srebrow and Kornblihtt 2006, Venables 2004, Venables et al 2009, Wang and Cooper 2007). Abnormal alternative splicing is caused by mutations in *cis*-acting elements including splicing enhance (ESE and ISE), and silencer (ESS and ISS). Furthermore

mutations or post-translational modification (phosphorylation, acetylation, etc) on the *trans*-acting factor RNA binding protein affect alternative splicing events. Therefore both *cis*-acting elements and *trans*-acting factors can change mRNA splicing profile.

Recent sequencing technology was significantly improved to apply investigation of splicing profile by whole transcripts sequencing, so called next-generation sequencing or highthough put sequencing (Marguerat et al 2008, Xiao and Lee 2010). Currently, 454 GS FLX Titanium (Roche) can read 400-600 million bp per single running and SOLiD system (Applied Biosystem) can generate 15 Giga bp with 50 nt long reads. Using this system the entire mRNA profile, in other words "transcriptome", can be monitored. Therefore, thousands of novel splicing variants by alternative splicing could be discovered. Moreover, cancer specific alternative splicing profile can be indentified by using combination of next-generation sequencing technique and systematic data analysis algorithm.

Another advanced technology for alternative splicing is a cross-linking and immunoprecipitation (CLIP) combined with high-throughput sequencing (HITS) technique (Licatalosi and Darnell 2010). CLIP method is usually applied for RNA-protein interaction study. Briefly, RNA-protein complex is covalently cross-linked by UV-irradiation, then RNA-protein complex is purified by immunoprecipitation. Next, protein bound regions on RNAs is sequenced using HITS technique. Therefore, the regulatory sequences including ESE, ESS, ISE, and ISS are identified in the whole transcriptome–wide by systematic sequence analysis. Those regulatory protein binding regions will represent the essential "footprint" to understand dynamic events of the alternative splicing regulation. However, there is limitation that has to be overcome the

specificity of antibody for the RNA binding protein. Finally, various newly developed techniques will provide global information to investigate the cancer specific alternative splicing.

CHAPTER 5

Conclusion:

Alternative splicing of KLF4 plays a role in tumorigenesis

Colorectal cancer is the third most common cancer and leading cause of cancer death in the United States. Colon cancer progression involves in the accumulation of sequential events that either activate oncogenes or inhibit the action of tumor suppressor genes. Particularly, Krüppel-like factor 4 (KLF4) is strongly associated with tumorigenesis and regulation of proliferation of GI tract epithelium (Wei et al 2006). In this dissertation research, we found novel KLF4 splicing variants produced by exon skipping in human cancer cell lines as well as colon tumor tissues (Figure 3-1 and Figure 3-2).

Most human genes including KLF4 produce different mRNAs by alternative splicing mechanism and many abnormal splicing processes are associated with human diseases, including cancer. However, the molecular relationship between alternative splicing and tumorigenesis is not well understood. Therefore, we hypothesized that alternative splicing of KLF4 is involved in colorectal cancer. To investigate mechanism of the KLF4 alternative splicing, we developed KLF4 minigene system (Figure 3-8). Mutation study of 3' splice sites on intron-2 showed that mutation of this site suppresses the KLF4 exon skipping (Figure 3-8). Moreover, sequence analysis revealed that the unique double 3' splice site contributes to in-frame alternative splice between $KLF4_L$ and $KLF4_S$ (Figure 3-9).

To identify *trans*-acting factor for KLF4 splicing, we performed gain and loss of functional studies with RBM5; RBM5 alters KLF4 splicing ratio (Figure 3-10). In addition, several anti-tumorigenic compounds were tested for KLF4 splicing (Figure3-11). Interestingly, sulindac sulfide restored $KLF4_L$ expression and this is mediated by dephosphorylation of RBM5 (Figure 3-12).

Furthermore, $KLF4_S$ localizes in the cytoplasm and nucleus and antagonizes $KLF4_L$ biological functions including the promoter binding activity to KLF4 target gene and binding with β-catenin (Figure 3-14, 15, 16, and 17).

Our data suggest that RBM5 plays a critical role in alternative splicing of KLF4 and $KLF4_S$ may contribute to colorectal tumorigenesis by competing with $KLF4_L$. In conclusion, we propose a model of the KLF4 alternative splicing mechanism (Figure 3-18). Finally, $KLF4_S$ can be used as a potential diagnostic and prognostic marker for human colorectal cancer.

REFERENCES

Abdel-Rahman WM, Peltomaki P (2004). Molecular basis and diagnostics of hereditary colorectal cancers. *Ann Med* **36:** 379-388.

Afify A, McNiel MA, Braggin J, Bailey H, Paulino AF (2008). Expression of CD44s, CD44v6, and hyaluronan across the spectrum of normal-hyperplasia-carcinoma in breast. *Appl Immunohistochem Mol Morphol* **16:** 121-127.

Alazzouzi H, Alhopuro P, Salovaara R, Sammalkorpi H, Jarvinen H, Mecklin JP *et al* (2005). SMAD4 as a prognostic marker in colorectal cancer. *Clin Cancer Res* **11:** 2606-2611.

Aoki K, Taketo MM (2007). Adenomatous polyposis coli (APC): a multi-functional tumor suppressor gene. *J Cell Sci* **120:** 3327-3335.

Asaoka Y, Tada M, Ikenoue T, Seto M, Imai M, Miyabayashi K *et al* (2010). Gastric cancer cell line Hs746T harbors a splice site mutation of c-Met causing juxtamembrane domain deletion. *Biochem Biophys Res Commun* **394:** 1042-1046.

Berget SM, Moore C, Sharp PA (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* **74:** 3171-3175.

Black DL (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72:** 291-336.

Bonnal S, Martinez C, Forch P, Bachi A, Wilm M, Valcarcel J (2008). RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. *Mol Cell* **32:** 81-95.

Boukis LA, Liu N, Furuyama S, Bruzik JP (2004). Ser/Arg-rich protein-mediated communication between U1 and U2 small nuclear ribonucleoprotein particles. *J Biol Chem* **279:** 29647-29653.

Brogna S, Wen J (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **16:** 107-113.

Burt R, Neklason DW (2005). Genetic testing for inherited colon cancer. *Gastroenterology* **128:** 1696-1716.

Cao W, Jamison SF, Garcia-Blanco MA (1997). Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA* **3:** 1456-1467.

Castro-Rivera E, Samudio I, Safe S (2001). Estrogen regulation of cyclin D1 gene expression in ZR-75 breast cancer cells involves multiple enhancer elements. *J Biol Chem* **276:** 30853-30861.

Ceulemans H, Bollen M (2004). Functional diversity of protein phosphatase-1, a cellular economizer and reset button. *Physiol Rev* **84:** 1-39.

Chan AT, Giovannucci EL, Meyerhardt JA, Schernhammer ES, Curhan GC, Fuchs CS (2005). Long-term use of aspirin and nonsteroidal anti-inflammatory drugs and risk of colorectal cancer. *JAMA* **294:** 914-923.

Chan AT, Giovannucci EL (2010). Primary prevention of colorectal cancer. *Gastroenterology* **138:** 2029-2043 e2010.

Chen YQ, Su M, Walia RR, Hao Q, Covington JW, Vaughan DE (1998). Sp1 sites mediate activation of the plasminogen activator inhibitor-1 promoter by glucose in vascular smooth muscle cells. *J Biol Chem* **273:** 8225-8231.

Cheng H, Dufu K, Lee CS, Hsu JL, Dias A, Reed R (2006). Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* **127:** 1389-1400.

Chow LT, Gelinas RE, Broker TR, Roberts RJ (1977a). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **12:** 1-8.

Chow LT, Roberts JM, Lewis JB, Broker TR (1977b). A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. *Cell* **11:** 819-836.

Cooper TA (2005). Use of minigene systems to dissect alternative splicing elements. *Methods* **37:** 331-340.

Coulter LR, Landree MA, Cooper TA (1997). Identification of a new class of exonic splicing enhancers by in vivo selection. *Mol Cell Biol* **17:** 2143-2150.

Dang DT, Bachman KE, Mahatan CS, Dang LH, Giardiello FM, Yang VW (2000a). Decreased expression of the gut-enriched Kruppel-like factor gene in intestinal adenomas of multiple intestinal neoplasia mice and in colonic adenomas of familial adenomatous polyposis patients. *FEBS Lett* **476:** 203-207.

Dang DT, Pevsner J, Yang VW (2000b). The biology of the mammalian Kruppel-like family of transcription factors. *Int J Biochem Cell Biol* **32:** 1103-1121.

David CJ, Chen M, Assanah M, Canoll P, Manley JL (2010). HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* **463:** 364-368.

Derynck R, Zhang YE (2003). Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425:** 577-584.

Downward J (2003). Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer* **3:** 11-22.

Duncan PI, Stojdl DF, Marius RM, Bell JC (1997). In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase. *Mol Cell Biol* **17:** 5996-6001.

Ekim M, Ikinciogullari A, Ulukol B, Bakkaloglu SA, Ozkaya N, Kendirli T *et al* (2003). Evaluation of nutritional status and factors related to malnutrition in children on CAPD. *Perit Dial Int* **23:** 557-562.

Eperon IC, Makarova OV, Mayeda A, Munroe SH, Caceres JF, Hayward DG *et al* (2000). Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Mol Cell Biol* **20:** 8303-8318.

Evans PM, Chen X, Zhang W, Liu C (2010). KLF4 interacts with beta-catenin/TCF4 and blocks p300/CBP recruitment by beta-catenin. *Mol Cell Biol* **30:** 372-381.

Fearon ER, Vogelstein B (1990). A genetic model for colorectal tumorigenesis. *Cell* **61:** 759-767.

Fearon ER (1995). Molecular genetics of colorectal cancer. *Ann N Y Acad Sci* **768:** 101-110.

Fedor MJ (2008). Alternative splicing minireview series: combinatorial control facilitates splicing regulation of gene expression and enhances genome diversity. *J Biol Chem* **283:** 1209-1210.

Foster KW, Frost AR, McKie-Bell P, Lin CY, Engler JA, Grizzle WE *et al* (2000). Increase of GKLF messenger RNA and protein expression during progression of breast cancer. *Cancer Res* **60:** 6488-6495.

Foster KW, Liu Z, Nail CD, Li X, Fitzgerald TJ, Bailey SK *et al* (2005). Induction of KLF4 in basal keratinocytes blocks the proliferation-differentiation switch and initiates squamous epithelial dysplasia. *Oncogene* **24:** 1491-1500.

Fushimi K, Ray P, Kar A, Wang L, Sutherland LC, Wu JY (2008). Up-regulation of the proapoptotic caspase 2 splicing isoform by a candidate tumor suppressor, RBM5. *Proc Natl Acad Sci U S A* **105:** 15708-15713.

Garcia-Blanco MA, Baraniak AP, Lasda EL (2004). Alternative splicing in disease and therapy. *Nat Biotechnol* **22:** 535-546.

Garrett-Sinha LA, Eberspaecher H, Seldin MF, de Crombrugghe B (1996). A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. *J Biol Chem* **271:** 31384-31390.

Ghigna C, Valacca C, Biamonti G (2008). Alternative splicing and tumor progression. *Curr Genomics* **9:** 556-570.

Goncalves V, Theisen P, Antunes O, Medeira A, Ramos JS, Jordan P *et al* (2009). A missense mutation in the APC tumor suppressor gene disrupts an ASF/SF2 splicing enhancer motif and causes pathogenic skipping of exon 14. *Mutat Res* **662:** 33-36.

Gorlach M, Burd CG, Portman DS, Dreyfuss G (1993). The hnRNP proteins. *Mol Biol Rep* **18:** 73-78.

Goss KH, Groden J (2000). Biology of the adenomatous polyposis coli tumor suppressor. *J Clin Oncol* **18:** 1967-1979.

Grady WM (2003). Genetic testing for high-risk colon cancer patients. *Gastroenterology* **124:** 1574-1594.

Grady WM (2005). Epigenetic events in the colorectum and in colon cancer. *Biochem Soc Trans* **33:** 684-688.

Graveley BR (2000). Sorting out the complexity of SR protein functions. *RNA* **6:** 1197-1211.

Grosso AR, Martins S, Carmo-Fonseca M (2008). The emerging role of splicing factors in cancer. *EMBO Rep* **9:** 1087-1093.

Guo J, Li Z-c, Feng Y-h (2009). Expression and activation of the reprogramming transcription factors. *Biochemical and Biophysical Research Communications* **390:** 1081-1086.

Guo TB, Boros LG, Chan KC, Hikim AP, Hudson AP, Swerdloff RS *et al* (2003). Spermatogenetic expression of RNA-binding motif protein 7, a protein that interacts with splicing factors. *J Androl* **24:** 204-214.

Han J, Cooper TA (2005). Identification of CELF splicing activation and repression domains in vivo. *Nucleic Acids Res* **33:** 2769-2780.

Hare LE, Ditzler CA, Hichens M, Rosegay A, Duggan DE (1977). Analysis of sulindac and metabolites by combined isotope dilution-radioimmunoassay. *J Pharm Sci* **66:** 414-417.

He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT *et al* (1998). Identification of c-MYC as a target of the APC pathway. *Science* **281:** 1509-1512.

Hofstetter G, Berger A, Fiegl H, Slade N, Zoric A, Holzer B *et al* (2010). Alternative splicing of p53 and p73: the novel p53 splice variant p53delta is an independent prognostic marker in ovarian cancer. *Oncogene* **29:** 1997-2004.

House AE, Lynch KW (2008). Regulation of alternative splicing: more than just the ABCs. *J Biol Chem* **283:** 1217-1221.

Huang CS, Shen CY, Wang HW, Wu PE, Cheng CW (2007). Increased expression of SRp40 affecting CD44 splicing is associated with the clinical outcome of lymph node metastasis in human breast cancer. *Clin Chim Acta* **384:** 69-74.

Jiang K, Patel NA, Watson JE, Apostolatos H, Kleiman E, Hanson O *et al* (2009). Akt2 regulation of Cdc2-like kinases (Clk/Sty), serine/arginine-rich (SR) protein phosphorylation, and insulin-induced alternative splicing of PKCbetaII messenger ribonucleic acid. *Endocrinology* **150:** 2087-2097.

Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD *et al* (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* **302:** 2141-2144.

Jurica MS, Moore MJ (2003). Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell* **12:** 5-14.

Kaczynski J, Cook T, Urrutia R (2003). Sp1- and Kruppel-like transcription factors. *Genome Biol* **4:** 206.

Kan Z, Garrett-Engele PW, Johnson JM, Castle JC (2005). Evolutionarily conserved and diverged alternative splicing events show different expression and functional profiles. *Nucleic Acids Res* **33:** 5659-5666.

Karni R, de Stanchina E, Lowe SW, Sinha R, Mu D, Krainer AR (2007). The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat Struct Mol Biol* **14:** 185-193.

Kataoka N, Yong J, Kim VN, Velazquez F, Perkinson RA, Wang F *et al* (2000). Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol Cell* **6:** 673-682.

Katz JP, Perreault N, Goldstein BG, Lee CS, Labosky PA, Yang VW *et al* (2002). The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. *Development* **129:** 2619-2628.

Katz JP, Perreault N, Goldstein BG, Actman L, McNally SR, Silberg DG *et al* (2005). Loss of Klf4 in mice causes altered proliferation and differentiation and precancerous changes in the adult stomach. *Gastroenterology* **128:** 935-945.

Kita H, Carmichael J, Swartz J, Muro S, Wyttenbach A, Matsubara K *et al* (2002). Modulation of polyglutamine-induced cell death by genes identified by expression profiling. *Hum Mol Genet* **11:** 2279-2287.

Klingbeil P, Marhaba R, Jung T, Kirmse R, Ludwig T, Zoller M (2009). CD44 variant isoforms promote metastasis formation by a tumor cell-matrix cross-talk that supports adhesion and apoptosis resistance. *Mol Cancer Res* **7:** 168-179.

Kotlajich MV, Hertel KJ (2008). Death by splicing: tumor suppressor RBM5 freezes splice-site pairing. *Mol Cell* **32:** 162-164.

Krecic AM, Swanson MS (1999). hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol* **11:** 363-371.

Kwiek NC, Thacker DF, Datto MB, Megosh HB, Haystead TA (2006). PITK, a PP1 targeting subunit that modulates the phosphorylation of the transcriptional regulator hnRNP K. *Cell Signal* **18:** 1769-1778.

Lai MC, Kuo HW, Chang WC, Tarn WY (2003). A novel splicing regulator shares a nuclear import pathway with SR proteins. *EMBO J* **22:** 1359-1369.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J *et al* (2001). Initial sequencing and analysis of the human genome. *Nature* **409:** 860-921.

Lee SH, Bahn JH, Choi CK, Whitlock NC, English AE, Safe S *et al* (2008). ESE-1/EGR-1 pathway plays a role in tolfenamic acid-induced apoptosis in colorectal cancer cells. *Mol Cancer Ther* **7:** 3739-3750.

Licatalosi DD, Darnell RB (2010). RNA processing and its regulation: global insights into biological networks. *Nat Rev Genet* **11:** 75-87.

Luk GD (1996). Prevention of gastrointestinal cancer--the potential role of NSAIDs in colorectal cancer. *Schweiz Med Wochenschr* **126:** 801-812.

Lynch HT, Smyrk TC, Watson P, Lanspa SJ, Lynch JF, Lynch PM *et al* (1993). Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. *Gastroenterology* **104:** 1535-1549.

Lynch HT (1999). Hereditary nonpolyposis colorectal cancer (HNPCC). *Cytogenet Cell Genet* **86:** 130-135.

Maquat LE, Carmichael GG (2001). Quality control of mRNA function. *Cell* **104:** 173-176.

Marguerat S, Wilhelm BT, Bahler J (2008). Next-generation sequencing: applications beyond genomes. *Biochem Soc Trans* **36:** 1091-1096.

Marnett LJ (1992). Aspirin and the potential role of prostaglandins in colon cancer. *Cancer Res* **52:** 5575-5589.

Martin A, Schneider S, Schwer B (2002). Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* **277:** 17743-17750.

Martinez-Contreras R, Cloutier P, Shkreta L, Fisette JF, Revil T, Chabot B (2007). hnRNP proteins and splicing control. *Adv Exp Med Biol* **623:** 123-147.

Masferrer JL, Zweifel BS, Manning PT, Hauser SD, Leahy KM, Smith WG *et al* (1994). Selective inhibition of inducible cyclooxygenase 2 in vivo is antiinflammatory and nonulcerogenic. *Proc Natl Acad Sci U S A* **91:** 3228-3232.

Massiello A, Chalfant CE (2006). SRp30a (ASF/SF2) regulates the alternative splicing of caspase-9 pre-mRNA and is required for ceramide-responsiveness. *J Lipid Res* **47:** 892-897.

Mathew R, Hartmuth K, Mohlmann S, Urlaub H, Ficner R, Luhrmann R (2008). Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nat Struct Mol Biol* **15:** 435-443.

Matlin AJ, Clark F, Smith CW (2005). Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* **6:** 386-398.

McGettigan P, Henry D (2006). Cardiovascular risk and inhibition of cyclooxygenase: a systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2. *JAMA* **296:** 1633-1644.

Mulder KM (2000). Role of Ras and Mapks in TGFbeta signaling. *Cytokine Growth Factor Rev* **11:** 23-35.

Naor D, Nedvetzki S, Golan I, Melnik L, Faitelson Y (2002). CD44 in cancer. *Crit Rev Clin Lab Sci* **39:** 527-579.

Narla G, Heath KE, Reeves HL, Li D, Giono LE, Kimmelman AC *et al* (2001). KLF6, a candidate tumor suppressor gene mutated in prostate cancer. *Science* **294:** 2563-2566.

Narla G, Difeo A, Reeves HL, Schaid DJ, Hirshfeld J, Hod E *et al* (2005). A germline DNA polymorphism enhances alternative splicing of the KLF6 tumor suppressor gene and is associated with increased prostate cancer risk. *Cancer Res* **65:** 1213-1222.

Ngo JC, Giang K, Chakrabarti S, Ma CT, Huynh N, Hagopian JC *et al* (2008). A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. *Mol Cell* **29:** 563-576.

Oh JJ, West AR, Fishbein MC, Slamon DJ (2002). A candidate tumor suppressor gene, H37, from the human lung cancer tumor suppressor locus 3p21.3. *Cancer Res* **62:** 3207-3213.

Pensabene M, Spagnoletti I, Capuano I, Condello C, Pepe S, Contegiacomo A *et al* (2009). Two mutations of BRCA2 gene at exon and splicing site in a woman who underwent oncogenetic counseling. *Ann Oncol* **20:** 874-878.

Piazza GA, Alberts DS, Hixson LJ, Paranka NS, Li H, Finn T *et al* (1997). Sulindac sulfone inhibits azoxymethane-induced colon carcinogenesis in rats without reducing prostaglandin levels. *Cancer Res* **57:** 2909-2915.

Pinto D, Gregorieff A, Begthel H, Clevers H (2003). Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes Dev* **17:** 1709-1713.

Reed R (1996). Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr Opin Genet Dev* **6:** 215-220.

Rennel ES, Varey AH, Churchill AJ, Wheatley ER, Stewart L, Mather S *et al* (2009). VEGF(121)b, a new member of the VEGF(xxx)b family of VEGF-A splice isoforms, inhibits neovascularisation and tumour growth in vivo. *Br J Cancer* **101:** 1183-1193.

Rintala-Maki ND, Goard CA, Langdon CE, Wall VE, Traulsen KE, Morin CD *et al* (2007). Expression of RBM5-related factors in primary breast tissue. *J Cell Biochem* **100:** 1440-1458.

Roesler J, Izquierdo JM, Ryser M, Rosen-Wolff A, Gahr M, Valcarcel J *et al* (2005). Haploinsufficiency, rather than the effect of an excessive production of soluble CD95 (CD95{Delta}TM), is the basis for ALPS Ia in a family with duplicated 3' splice site AG in CD95 intron 5 on one allele. *Blood* **106:** 1652-1659.

Roscigno RF, Garcia-Blanco MA (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* **1:** 692-706.

Rowland BD, Peeper DS (2006). KLF4, p21 and context-dependent opposing forces in cancer. *Nat Rev Cancer* **6:** 11-23.

Schreibman IR, Baker M, Amos C, McGarrity TJ (2005). The hamartomatous polyposis syndromes: a clinical and molecular review. *Am J Gastroenterol* **100:** 476-490.

Segre JA, Bauer C, Fuchs E (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat Genet* **22:** 356-360.

Shen TY, Winter CA (1977). Chemical and biological studies on indomethacin, sulindac and their analogs. *Adv Drug Res* **12:** 90-245.

Shie JL, Chen ZY, Fu M, Pestell RG, Tseng CC (2000). Gut-enriched Kruppel-like factor represses cyclin D1 promoter activity through Sp1 motif. *Nucleic Acids Res* **28:** 2969-2976.

Shields JM, Yang VW (1998). Identification of the DNA sequence that interacts with the gut-enriched Kruppel-like factor. *Nucleic Acids Res* **26:** 796-802.

Sieber OM, Tomlinson IP, Lamlum H (2000). The adenomatous polyposis coli (APC) tumour suppressor--genetics, function and disease. *Mol Med Today* **6:** 462-469.

Singh A, Karnoub AE, Palmby TR, Lengyel E, Sondek J, Der CJ (2004). Rac1b, a tumor associated, constitutively active Rac1 splice variant, promotes cellular transformation. *Oncogene* **23:** 9369-9380.

Sogawa K, Imataka H, Yamasaki Y, Kusume H, Abe H, Fujii-Kuriyama Y (1993). cDNA cloning and transcriptional properties of a novel GC box-binding protein, BTEB2. *Nucleic Acids Res* **21:** 1527-1532.

Solomon SD, McMurray JJ, Pfeffer MA, Wittes J, Fowler R, Finn P *et al* (2005). Cardiovascular risk associated with celecoxib in a clinical trial for colorectal adenoma prevention. *N Engl J Med* **352:** 1071-1080.

Sorek R, Shamir R, Ast G (2004). How prevalent is functional alternative splicing in the human genome? *Trends Genet* **20:** 68-71.

Srebrow A, Kornblihtt AR (2006). The connection between splicing and cancer. *J Cell Sci* **119:** 2635-2641.

Sukhthankar M, Yamaguchi K, Lee SH, McEntee MF, Eling TE, Hara Y *et al* (2008). A green tea component suppresses posttranslational expression of basic fibroblast growth factor in colorectal cancer. *Gastroenterology* **134:** 1972-1980.

Suske G, Bruford E, Philipsen S (2005). Mammalian SP/KLF transcription factors: bring in the family. *Genomics* **85:** 551-556.

Sutherland LC, Rintala-Maki ND, White RD, Morin CD (2005). RNA binding motif (RBM) proteins: a novel family of apoptosis modulators? *J Cell Biochem* **94:** 5-24.

Tarn WY, Steitz JA (1996). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* **84:** 801-811.

Tetsu O, McCormick F (1999). Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* **398:** 422-426.

Thun MJ, Namboodiri MM, Heath CW, Jr. (1991). Aspirin use and reduced risk of fatal colon cancer. *N Engl J Med* **325:** 1593-1596.

Ton-That H, Kaestner KH, Shields JM, Mahatanankoon CS, Yang VW (1997). Expression of the gut-enriched Kruppel-like factor gene during development and intestinal tumorigenesis. *FEBS Lett* **419:** 239-243.

Turner J, Crossley M (1999). Mammalian Kruppel-like transcription factors: more than just a pretty finger. *Trends Biochem Sci* **24:** 236-240.

Ulrich CM, Bigler J, Potter JD (2006). Non-steroidal anti-inflammatory drugs for cancer prevention: promise, perils and pharmacogenetics. *Nat Rev Cancer* **6:** 130-140.

Vane JR (1971). Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nat New Biol* **231:** 232-235.

Venables JP (2004). Aberrant and alternative splicing in cancer. *Cancer Res* **64:** 7647-7654.

Venables JP, Klinck R, Koh C, Gervais-Bird J, Bramard A, Inkel L *et al* (2009). Cancer-associated regulation of alternative splicing. *Nat Struct Mol Biol* **16:** 670-676.

Walther A, Johnstone E, Swanton C, Midgley R, Tomlinson I, Kerr D (2009). Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer* **9:** 489-499.

Wang GS, Cooper TA (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* **8:** 749-761.

Wang J, Yu P, Lin TC, Konigsberg WH, Steitz TA (1996). Crystal structures of an NH2-terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions. *Biochemistry* **35:** 8110-8119.

Wei D, Gong W, Kanai M, Schlunk C, Wang L, Yao JC *et al* (2005). Drastic down-regulation of Kruppel-like factor 4 expression is critical in human gastric cancer development and progression. *Cancer Res* **65:** 2746-2754.

Wei D, Kanai M, Huang S, Xie K (2006). Emerging role of KLF4 in human gastrointestinal cancer. *Carcinogenesis* **27:** 23-31.

Wielenga VJ, Smits R, Korinek V, Smit L, Kielman M, Fodde R *et al* (1999). Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. *Am J Pathol* **154:** 515-523.

Williams CS, Goldman AP, Sheng H, Morrow JD, DuBois RN (1999). Sulindac sulfide, but not sulindac sulfone, inhibits colorectal cancer growth. *Neoplasia* **1:** 170-176.

Wu JY, Maniatis T (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75:** 1061-1070.

Xiao SH, Manley JL (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev* **11:** 334-344.

Xiao X, Lee JH (2010). Systems analysis of alternative splicing and its regulation. *Wiley Interdiscip Rev Syst Biol Med* **2:** 550-565.

Yamaguchi K, Lee SH, Eling TE, Baek SJ (2006). A novel peroxisome proliferator-activated receptor gamma ligand, MCC-555, induces apoptosis via posttranscriptional regulation of NAG-1 in colorectal cancer cells. *Mol Cancer Ther* **5:** 1352-1361.

Yoon HS, Chen X, Yang VW (2003). Kruppel-like factor 4 mediates p53-dependent G1/S cell cycle arrest in response to DNA damage. *J Biol Chem* **278:** 2101-2105.

Zhang W, Geiman DE, Shields JM, Dang DT, Mahatan CS, Kaestner KH *et al* (2000). The gut-enriched Kruppel-like factor (Kruppel-like factor 4) mediates the transactivating effect of p53 on the p21WAF1/Cip1 promoter. *J Biol Chem* **275:** 18391-18398.

Zhang W, Chen X, Kato Y, Evans PM, Yuan S, Yang J *et al* (2006). Novel cross talk of Kruppel-like factor 4 and beta-catenin regulates normal intestinal homeostasis and tumor repression. *Mol Cell Biol* **26:** 2055-2064.

Zhao W, Hisamuddin IM, Nandan MO, Babbin BA, Lamb NE, Yang VW (2004). Identification of Kruppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer. *Oncogene* **23:** 395-402.

Zhong XY, Ding JH, Adams JA, Ghosh G, Fu XD (2009). Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes Dev* **23:** 482-495.

VITA

Jae Hoon Bahn was born Nov 11, 1973 in Seoul, South Korea. He received the degree of Bachelor of Science from Hallym University in 1999. In the middle of undergraduate study, he served in Korean Army for 29 months. After graduation, he continued his study with Dr. Soo Young Choi in the graduate school at Hallym University. He received the Master of Science degree in 2001. He started graduate study at the University of Tennessee, Knoxville in 2003. He joined Dr. Seung Joon Baek' Laboratory in 2007. Since undergraduate, he has 38 peer-reviewed publications including *Oncogene, PNAS,* and *Genetics*. He graduated with a Doctor of Philosophy in Comparative and Experimental Medicine in the December 2010. He is beginning his postdoctoral training to study mechanisms of alternative splicing at the University of California, Los Angeles.

Publications

1. **Bahn JH**, Lee SH, Yoon JH, and Baek SJ. Alternative splicing of Kruppel-like factor 4 plays a role in colorectal tumorigenesis. ***Oncogene*** 2010 (Under revision)

2. Whitlock NC, **Bahn JH,** Lee SH, Eling TE, and Baek SJ. Early Growth Response-1 and Krupel-like Factor 4 regulate Resveratrol-induced Activating Transcription Factor 3 expression. ***Cancer prevention research*** 2010 (In Press)

3. Flatland B, Fry MM, Baek SJ, **Bahn JH**, LeBlanc CJ, Dunlap JR, Carroll RC, Kosiba DJ, Millsaps DJ, Schleis SE. May-hegglin anomaly in a dog. ***Vet Clinical Patho*** 2010 (In press)

4. Kim CH, **Bahn JH**, Lee SH, Kim GY, Jun SI, Lee K, and Baek SJ. Induction of cell growth arrest by atmospheric non-thermal plasma in colorectal cancer cells. ***J Biotechnol*** 2010 (In press)

5. Lee SH, **Bahn JH**, Whitlock NC, and Baek SJ. Activating transcription factor 2 (ATF2) controls tolfenamic acid-induced ATF3 expression via MAP kinase pathways. ***Oncogene*** 2010, 29; 5182-5192.

6. Kim CH, Kwon S, **Bahn JH**, Lee K, Jun SI, Rack PD and Baek SJ. Effects of atmospheric non-thermal plasma on invasion of colorectal cancer cells. ***Applied Physics Letters*** 2010, 96; 243701.

7. **Bahn JH**, Lee G, and Park JH. Comparative analysis of Pdf-mediated circadian behaviors between *Drosophila melanogaster* and *D. virilis*. ***Genetics*** 2009, 181; 965-75.

8. Lee SH, **Bahn JH**, Choi CK, Whitlock NC, English AE, Safe S, and Baek SJ. ESE-1/EGR-1 pathway plays a role in tolfenamic acid-induced apoptosis in colorectal cancer cells. ***Mol Cancer Ther***. 2008, 7; 3739-50.

9. Lee G, **Bahn JH**, and Park JH. Sex- and clock-controlled expression of the neuropeptide F gene in Drosophila. ***Proc Natl Acad Sci U S A***. 2006, 103; 12580-12585.